

Geographic Information System Integration with the “A New Nation Votes” Project

Project: GIS Historic Voting Patterns

IQP Number: LES 2654

Author: Kristopher Kellogg

Faculty Advisor: Professor Lance Schachterle

Abstract: This project, in association with the American Antiquarian Society, describes the procedure involved with generating graphical displays of the Phillip Lampi electoral data collection. The required data were taken from the “A New Nation Votes” website in order to generate maps using the geographic information system ArcGIS. In addition, ArcGIS was used to transform existing historical maps into a digital format, making the digitized electoral data and historical maps more user-friendly and useful for further analysis.

Acknowledgements:

Professor Paul Mathisen for providing a general tutorial for the ArcGIS software package.

Erik Beck at the American Antiquarian Society for providing input and guidance for the project.

Introduction:

This project, in association with the American Antiquarian Society, describes the procedure involved with generating graphical displays of the Phillip Lampi electoral data collection. The required data were taken from the "A New Nation Votes" website in order to generate maps using the geographic information system ArcGIS. This will further the use of the digitized electoral data collection by enabling the creation of a user-friendly format to view that data.

GIS stands for geographic information system. This is a system that is used to display data in a spatially oriented setting, such as the layers of a map. This system can be used to show and manipulate any form of data that is specific to a location, including things like natural resources, the path of a spreading disease, or census data. By spatially representing these data, GIS enables quick analysis to occur that may have been impossible in another format.

The data used for GIS is stored in databases or separate files and can then be projected over an existing graph. In this way, GIS software links geographical data by spatial coordinates as well as time and then the user can manipulate the display of the data. This can be accomplished by changing which layers are displayed on top of other layers. An example of this would be a graph that displayed a state, the population density of the state, and highways sorted by use. In order to make a graph that shows these layers effectively, one would probably put the highways on the top, followed by the population density layer, and the state borders on the bottom layer. This enables the more specific data to be seen first, over broader areas of data.

The applications for GIS vary greatly, as nearly any use of information specific to a region could be improved using GIS software. Some major uses fall into the categories of logistics, urban and development planning, hydrology, and crime mapping. These fields all greatly benefit from the ability to coordinate data spatially and temporally as opposed to in a spreadsheet. Criminology, for instance, could follow the path of a criminal to narrow a search area down and increase the chance of catching

the person before any other crimes are committed. Hydrology can follow the trend of a river or stream and better prepare people around it for flooding. As shown, the analysis of data is simplified via GIS and can lead to more timely and accurate responses.

For my specific project GIS was used to represent voting records over a map that would range from the county level to all of the states and territories of 1825. These data could be used to generate maps in order to represent trends in the data, such as how a state voted over the course of several elections. This specific application had difficulties, however, due to the fluid nature of county and state lines during this period. The main objective with this application was to produce maps that are easy to read for the user and provide trend data as well as data for specific space-time points. These data would show how a country or state voted, perhaps how much of a change that was from a previous election, or how far a candidate was ahead of his opponent.

The data used for this project came from the Phillip Lampi collection of voting records available on the "A New Nation Votes" website. Mr. Lampi has been collecting electoral data for any kind of election within the United States from 1787 to 1825 for the past 38 years. This involved going to libraries, city halls, and archives all around the eastern United States in order to find data to copy. Much of his collection process took place by hand because it was done before photocopiers were prevalent. Every office from town clerk to President was taken into account. Every election from every year was important and had to be tabulated. This quest for information eventually resulted in a position with the American Antiquarian Society where Mr. Lampi currently is in the process of digitizing his data as well as collecting even further elections to add to the data (<http://h2g2.com>). Upon completion, the election data will include all 25 states and territories that existed in 1825. The goal of Tufts University and the American Antiquarian Society for these data is to make it available to all for free and to enable a fresh look at the electoral process within the United States when the nation was new (http://elections.lib.tufts.edu/aas_portal/about.html).

The software package used for this project was ArcGIS. This software package enables the user to make maps, display geographic data from a database, and analyze that data. For the scope of this project, ArcGIS enables voting records from the Lampi collection to be displayed geographically over a map of the states or counties relevant to the election. In order to display the results a map for the specific time period must exist, however, and this involved editing shape files for ArcGIS. Shape files in this case define the borders of the state or territory and the counties within. Some maps for the states and their counties are available through the Newberry Library website; others had to be created using coordinates or an approximation of coordinates. These maps show borders for each state and county or even proposed counties from the creation of the state until 2000. In order to create a map for a specific area and date, the first step was to download the appropriate GIS zip file.

For this example, a map of Massachusetts on May 1st 1800 was used. Once the GIS zip file for Massachusetts had been downloaded and extracted, ArcMap was used to specify the conditions for our map. This involved using the Add Data function in ArcMap to add the state's master polygon shape file, which has all the borders for the state that have existed, and using ArcMap's query builder to specify the date of the map in the following format: "START_N" <= 18000501 AND "END_N" >= 18000501 . This line within the define query box will tell the program that you wish to view the map of May 1st 1800. In order to change the date, one simply has to change 18000501 to the date desired in yyyyymmdd format. This specified map was then saved as a shape file (which would be used as its own map file) or as a feature class (which would be used to modify an existing geodatabase).

Once the borders of the base map were completed, it was ready for the data to be added. The data being used came from the Tufts digital election library. Elections were searched through on their website and data for states or the whole collection can be downloaded and viewed as a spreadsheet. The data translating step was the focus of this project: how could the data from the Lampi collection be added to an ArcGIS map in order to be used to create useful displays of election data? First the data had

to be put into a format that ArcGIS can read, such as coordinates. Then, once the data could be read by the program, it must be sorted in a way that makes sense. This involved moving layers of the map in order to display the data in an intuitive way. Finally, shading of layers and other convenient display features such as border thickness were used to generate maps of the data in a way that was clear to see for the user and showed the relevant information in an obvious way.

Literature Review:

Although I was unable to find any projects that were very similar to the scope of this project, there were many examples available about the uses of GIS. For this project, the intention is to inform users via a simplistic comparative display of multiple maps. This involves transforming pure data from a spreadsheet into a visual representation via ArcGIS that can be easily understood by someone with no knowledge of the software. Many other projects have tackled similar issues of displaying information for a layperson in order to enable them to draw conclusions from the data.

One example of such a task can be seen in Geographic Information Systems for Group Decision Making: Towards a Participatory, Geographic Information Science. In this book the authors discuss GIS technology being used in a group environment in order to benefit the decision making process by better informing those involved. If the data were scattered in several sources, or even organized in a fashion that is less visually understandable, such as a database, the decision making process of the group might be hindered. The use of GIS software enables the data to be easily viewed and analyzed by participants.

A similar article, "Internet GIS for Public Participation", shows how an internet-based GIS program can be used to display information across spatial and temporal lines in order to get the public more involved in the decision making process for things like community planning. This same process could be used for any sort of group decision making, such as transportation, public services, or

neighborhood development. These two articles both show how GIS has been used to display data in an easy to use format for the general public.

Alternatively, there were several articles that dealt with the voting process and GIS. These articles relate to the topic of this project because they deal with mapping and analyzing voting records, although they are more current than the Lampi data set. One such article is “Analyzing Partisanship in Central Mexico: A Geographical Approach.” In this article, GIS software is used to analyze trends in voting due to economic and educational differences. Another similar article is “Location, Knowledge and Time Pressures in the Spatial Structure of Convenience Voting” which looks at how big of a role convenience is when it comes to voting. For instance, people who are closer to voting areas may be more likely to vote because it is more convenient, whereas people with a long daily commute might be less likely due to time constraints.

Other articles involving GIS were of interest, although not directly related to the subject matter of this project. These articles were used to show the state of the art when it comes to geographic information systems. One example of this is an article entitled “A GIS Based Assessment of Potential for Windfarms in India.” In this article, the authors show how GIS was used to create a grid which used data about wind speeds in the country in order to determine the optimal locations for windfarms and to calculate the total amount of energy India could produce from windfarms. These calculations are obviously not reasonable values for energy production because they would involve covering most of the country in windfarms, but they were useful for studying potential windfarm sites.

Another article that could be seen as the cutting edge for GIS technology is “Arsenic Levels in the Soil and Risk of Birth Defects: A Population-Based Case-Control Study Using GIS Technology.” This article is very current, as it was printed in 2011, and it discusses a study which uses GIS technology in order to map levels of arsenic in the soil versus birth defects within a region. These correlations can be used to show how arsenic in the soil is directly linked to such birth defects. Such a study could be used

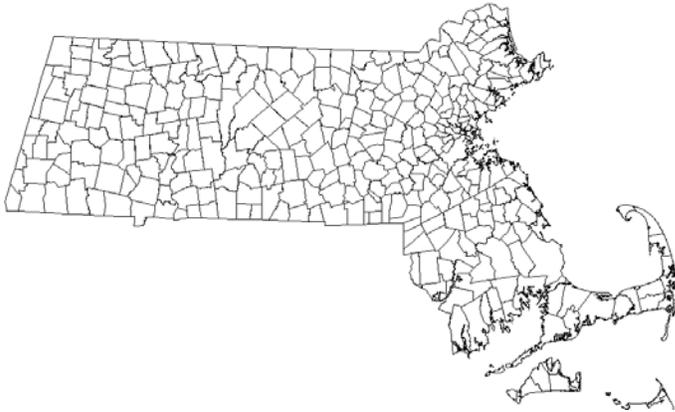
to map almost any disease or illness and its spread, enabling people to know where the disease originated from and where it is likely to spread, as well as how severe the risk is within a certain area.

Procedure:

I was asked by Erik Beck, the project coordinator for “A New Nation Votes” from the AAS, to use ArcGIS to create a comparison between the towns of Roxbury and Dorchester in 1800 and the current neighborhoods of Roxbury and Dorchester within Boston using a U. S. Congressional election from 1800. In order to accomplish this task several maps were needed. Maps of Dorchester and Roxbury from 1800 and today were necessary to show how the areas have changed. Also, a map of the surrounding area was necessary as the election chosen includes “District Middle 1”, which includes Boston and the surrounding area. To satisfy this requirement a map of Massachusetts from 1800 was required. Finally, a current map of Massachusetts was used in order to provide a comparison for the area of Boston and the surrounding counties. Once all of these maps were acquired it was possible to create an effective comparison of the two areas.

A GIS file for a current map of Massachusetts was easily found on the state’s GIS website, <http://www.mass.gov/mgis/massgis.htm>. This information could simply be loaded into ArcGIS using the Add Data function. These data could be manipulated within ArcGIS, as can all map files that one can

add, in order to provide the desired display.



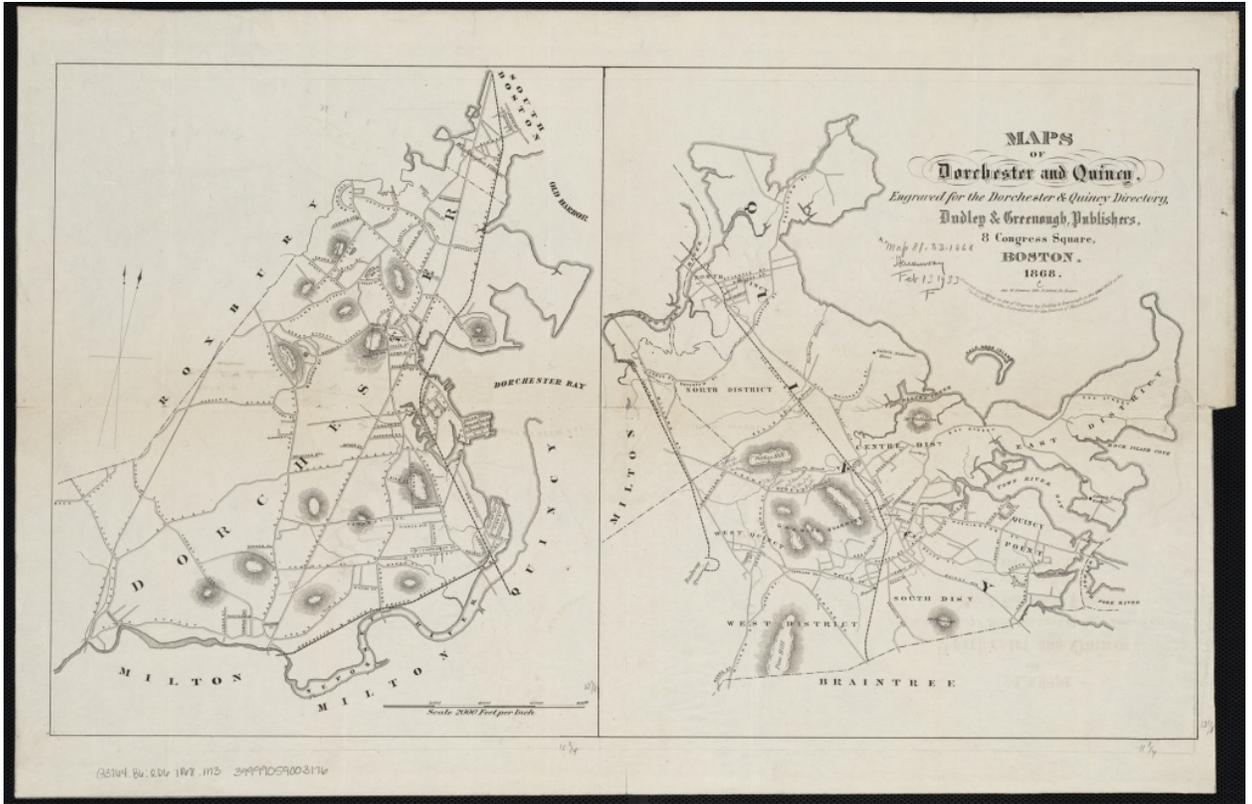
The Massachusetts map used is divided into towns, as shown above, with many features (each town) that can be selected and manipulated. Because I wished to show only Boston and the surrounding areas, I selected features for Middlesex, Norfolk and Suffolk counties and created a layer file from those features using the Export All Features option. This layer file can then be added to ArcGIS using the Add Data function at any time and then positioned, colored, and further manipulated to create the desired map. In this way, some features from a larger data set can be isolated to meet the needs of a given task, enabling the user to create a map with only the desired features.

For the 1800 map of Massachusetts, GIS files were obtained from the Newberry Library's website, <http://publications.newberry.org/ahcbp/>. This information was almost as simple to use as the current map of Massachusetts, as the GIS file package could be added in the same way. However, once the data was added it was necessary to manipulate the map in order to show the correct date. This was done by going into the data layer properties and adding a query with the Query Builder function. The line "START_N" <= yyyymmdd AND "END_N" >= yyyymmdd was used to specify the date of the map desired in year, month, day format. The data included in this package will show a map for the date of the creation of the first county within a given state up through December 31st 2000 when this query is provided. Thus, an input of 18000101 for both the start and end dates was used to specify a date of

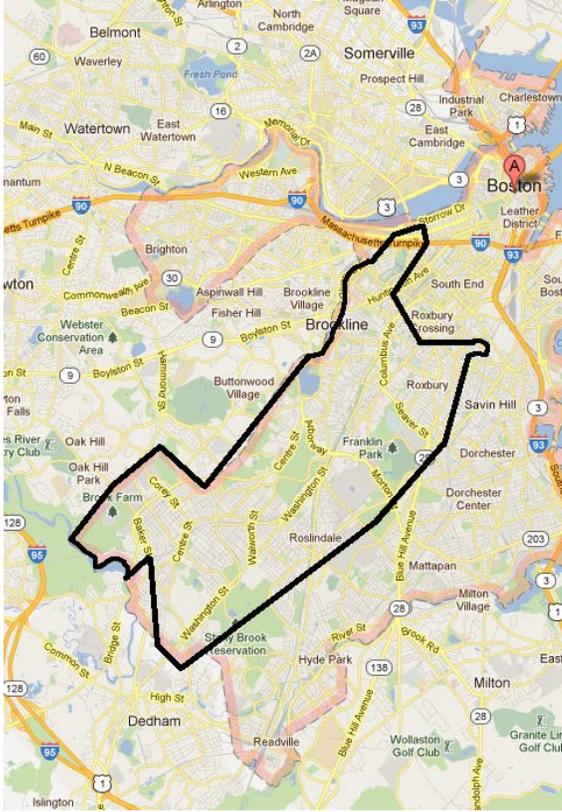
January 1st 1800. Once specified, the user will notice the default map change to meet the specified settings. Features of this map can then be selected and manipulated as described above in order to create a layer file.

Obtaining the remainder of the maps required to create the requested comparison was much more challenging. Map files for Roxbury and Dorchester in 1800 or the present could not be located when I looked for them on any Massachusetts government or historical society websites. An ArcGIS map of circa 1800 Roxbury or Dorchester is most likely too specific of a request for the Massachusetts government to have generated at this point, as GIS technology is relatively new and more current maps are likely in greater demand. Furthermore, because those towns are now neighborhoods it was extremely difficult to find a map file that broke Boston down into anything smaller than the city itself because neighborhoods are not a politically relevant division. Thus, a non-official website (<http://www.zillow.com/howto/api/neighborhood-boundaries.htm>) had to be used to obtain neighborhood maps and I had to generate my own maps of the 1800 areas.

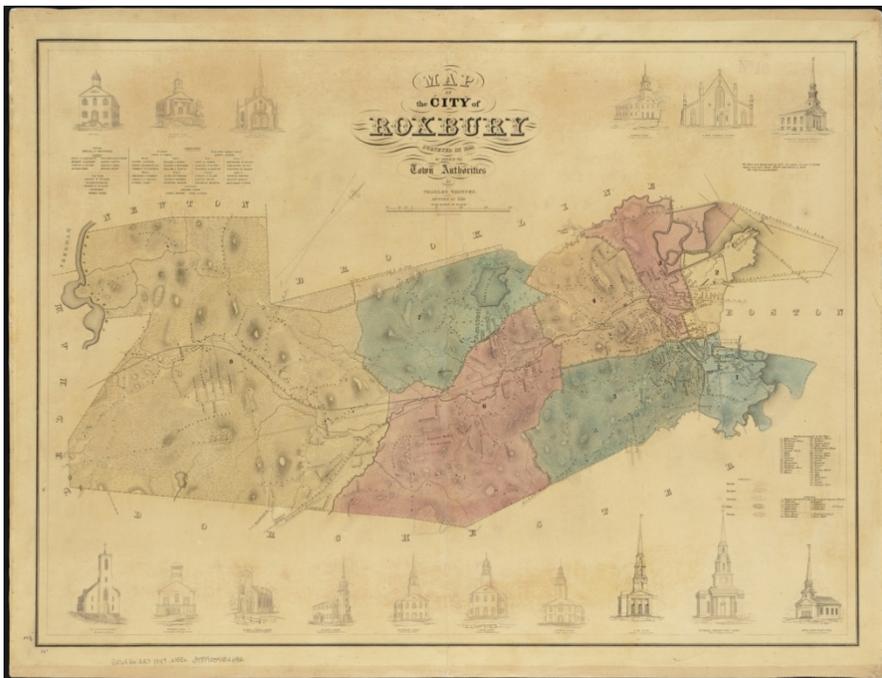
In order to generate layer files for 1800 Dorchester and Roxbury I had to obtain coordinate information for the maps in some way. I attempted to find GPS style or latitude and longitude coordinates for these areas and was unable to because the historical maps from this period didn't contain any of this information. With this information I could have created a spreadsheet of coordinates and imported the information into ArcGIS using the Add XY Data feature, which would have given a series of points which could then be connected using the polyline function. However, because I was unable to find any exact coordinate information I had to estimate using maps from the Boston Public Library found at <http://maps.bpl.org/>. Maps are only available for each town from select dates, so I used Historical Data Relating to Counties, Cities and Towns in Massachusetts to determine when the political borders of each town had changed in order to see which map would most closely match the



Historical Map of Dorchester 1868 (Left)



City of Roxbury 1849, comparison map.



Historical Map of Roxbury 1849

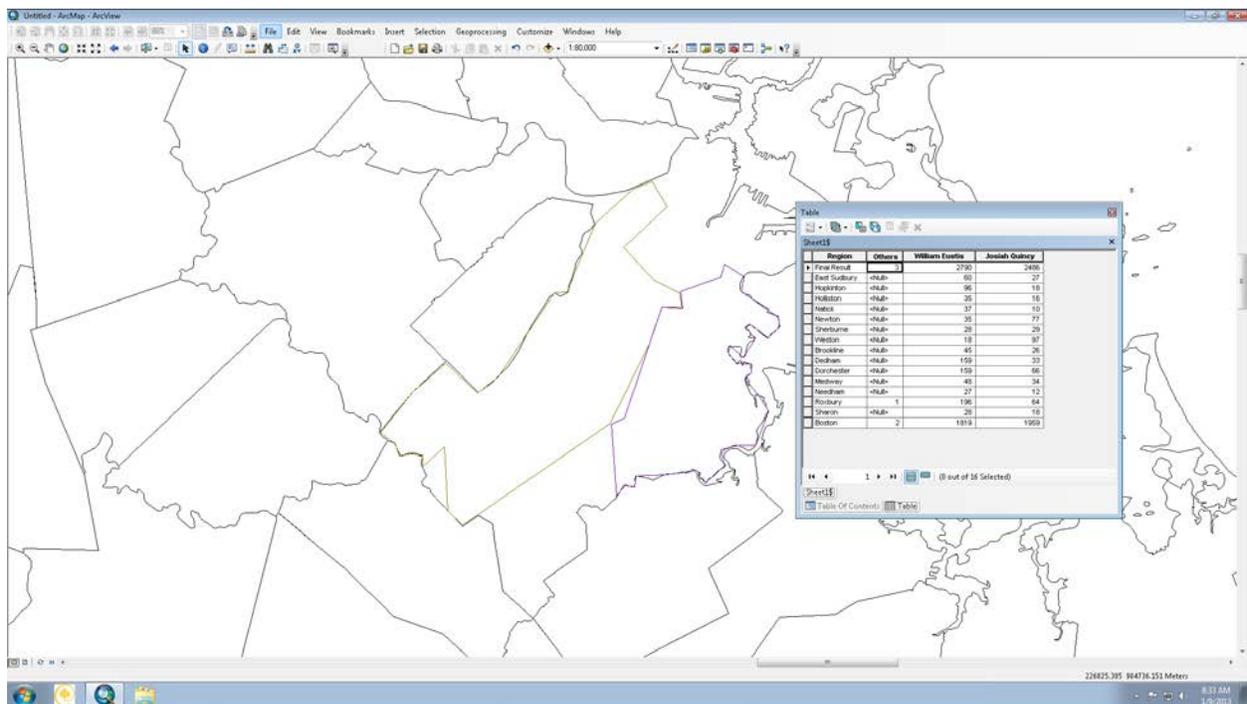
In order to generate a layer file for these historic borders the points I had used with Google maps had to be transferred into ArcGIS in some fashion. I accomplished this by loading a map of present day Massachusetts (the same map discussed above) and isolating the Boston feature, providing a side by side comparison to work with. Then, by utilizing the coordinate system displayed on the bottom right portion of the screen within ArcGIS, I began to obtain points by pointing my cursor to a border point and putting the coordinate displayed into a spreadsheet. Once sufficient points were obtained in this way to generate the shape I desired the border to resemble, the spreadsheet data was imported into ArcGIS and expressed as points on the map of Massachusetts. These points were then connected using the polyline function and saved as a layer file. A screenshot of these two areas is shown below with Roxbury in brown and Dorchester in purple.



This method of creating a layer file from estimated points leaves a lot of room for error, and so the layers do not perfectly represent the borders of the towns as they were in 1800. However, with the very limited information available this method seemed to be the only viable option for creating a map of the specified area. The process of obtaining and transferring points was painstaking, time-consuming

and somewhat inaccurate despite careful estimation. Thus, if at all possible this procedure should be avoided if coordinate data is available.

The voting records for the specified election were manually transferred into a spreadsheet file and added to ArcGIS using the add data function. Once added, this information was used to generate a table which could be displayed side by side with the map if desired by right clicking on the spreadsheet in the table of contents and choosing to open it. This would give the viewer all of the information about the candidates and the number of votes each obtained from a given town along with the visual display of the map. An example of this is shown below:



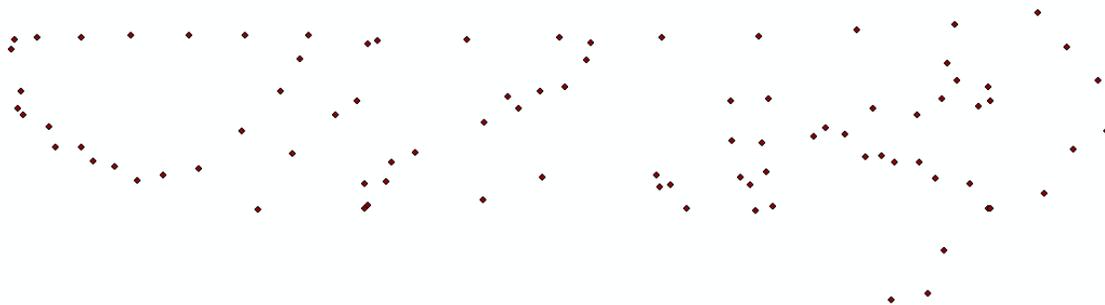
This display is not much more useful than a spreadsheet in an outside program would be other than the fact that it can be shown from within the program simultaneously with the map features.

Alternatively, this election information was used to alter the layer files in order to more simply display the records. This was accomplished by selecting the features that a candidate won (towns in this case) and creating a new layer file with just these features. In this way, for instance, a two candidate election could easily be sorted so as to display one candidate in one color and the other in a different

color, as modern Presidential elections are analyzed with red and blue states. For this particular election this method is not very effective, as only the towns of Roxbury and Dorchester are represented out of the list of towns and both towns were won by the same candidate. However, the same procedure can be applied to many elections in order to produce a simple display of election results.

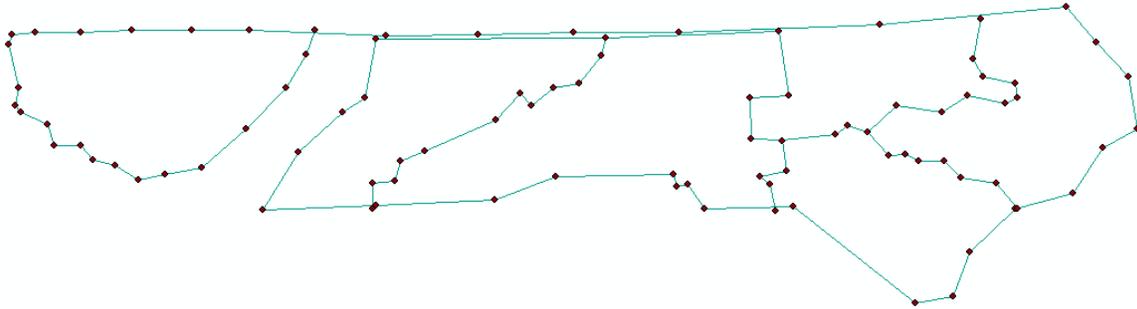
Erik Beck also asked me to generate maps for North Carolina from 1789 to 1825 in a similar fashion to those of Dorchester and Roxbury which were already discussed. The procedure for these maps was slightly different, in that the source material was found in [The Historical Atlas of United States Congressional Districts 1789-1983](#) instead of from digital versions of historical maps. In order to create layer files in ArcGIS from these images, I decided to convert them into an image file which would enable them to be used for direct translation of plot points.

These plot points were obtained by moving my cursor along the area where the district lines would be in the congressional district maps in the ArcGIS interface. ArcGIS provides a coordinate point for the location of the cursor, which enabled me to convert points on the image file for the map into a coordinate. Each coordinate point was then input into a spreadsheet and imported into ArcGIS, and the result of one example is shown below:

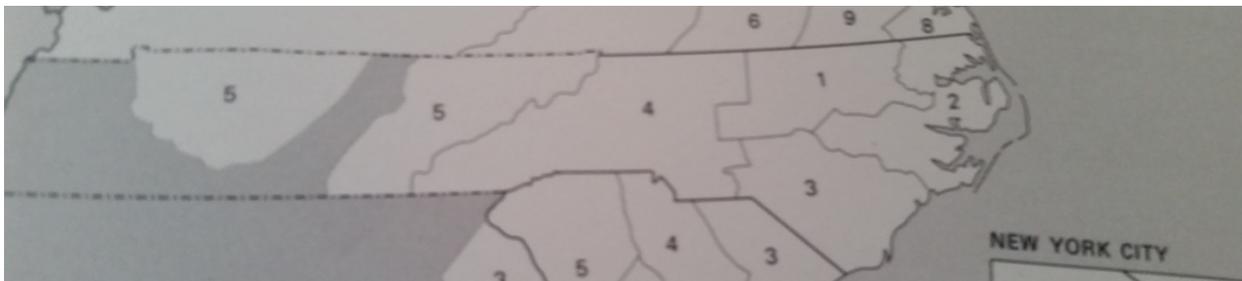


These points were then connected with the polyline function. This function connects points in the order they are entered with a solid line, so it is important to input the plot points in such a way that they will connect to form the outline desired. Because maps contain several district lines as well as state borders,

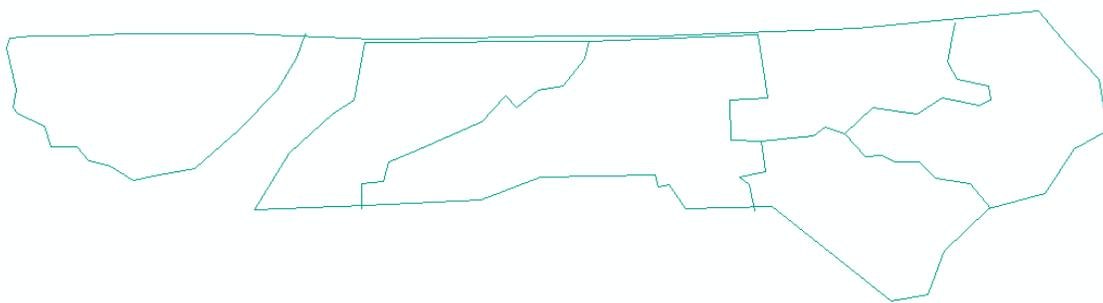
some overlap will occur in order to connect all line segments. An example of the connected points is seen here:



The connected points form the shape of the districts within the state as well as the state borders themselves. This particular example is modeled on the image seen below:

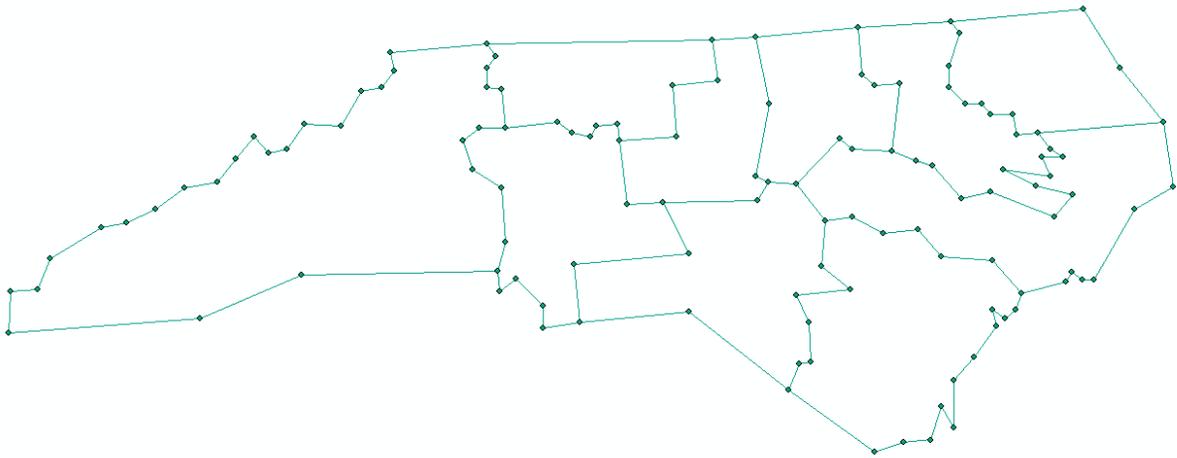


And when the plot points are hidden, the finished layer file will look like this:

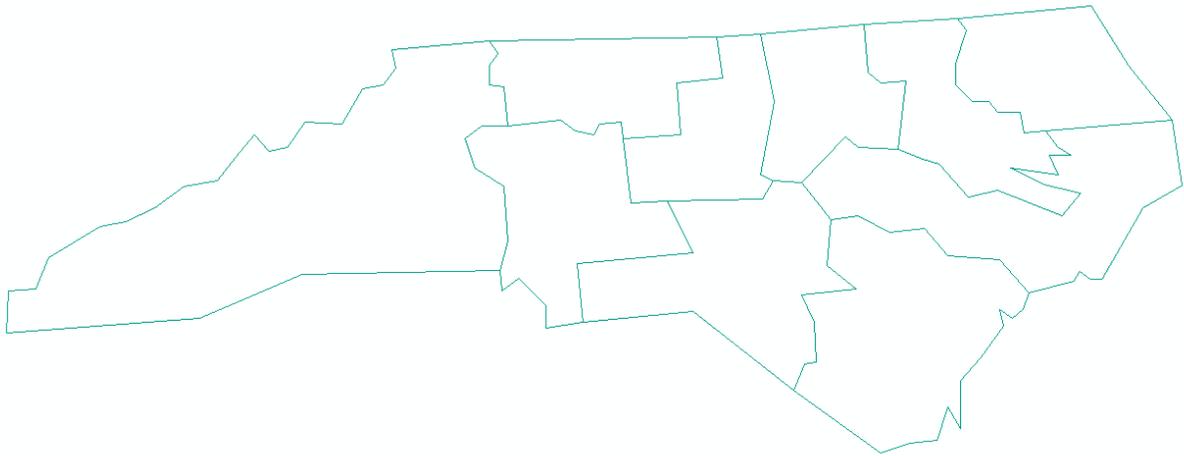


This layer file can be used to tie information, such as voting records and election results, to the digital representation of what the atlas displays.

The district line maps involved many more plot points than the Dorchester and Roxbury examples that were discussed earlier because every line segment requires plot points in order to be displayed. This means that as the number of districts within the state increases, so does the complexity of creating a layer file. Between the 1st Congress (the North Carolina example shown above is the 1st Congress from 1789-1791) and the 3rd Congress, the number of districts in the state of North Carolina doubled from 5 to 10. To maintain the same level of detail, this dictates that the number of plot points must also increase in order to reflect the new district shapes within the state. This is displayed in the maps below:



The 3rd Congress of North Carolina with plot points displayed



The 3rd Congress of North Carolina without plot points

These layer files can be used along with the historical voting records collected in the “A New Nation Votes” project to create an interactive, digital representation of data that was at one time only accessible in analog form. This will enable a much wider audience to access and explore the available information. In this way, ArcGIS is providing an outlet for Phillip Lampi’s collection of voting records gathered over a 38 year period to be shared with an increasingly digital generation.

Bibliography

- Aguirre, Robert, and Timothy Nyerges. "Geovisual Evaluation of public participation in decision making: The Grapevine." *Journal of Visual Languages and Computing* 22 (2011): 305-21. Print.
- "Atlas of Historical County Boundaries Project | Home." *Newberry Library Publications*. The Newberry Library, July 2010. Web. 12 Dec. 2011. <<http://publications.newberry.org/ahcbp/>>.
- Bertelli, Anthony M., and Jamie L. Carson. "Small Changes, Big Results: Legislative Voting Behavior in the Presence of New Voters." *Electoral Studies* 30.1 (2011): 201-09. Print.
- Chang, Grace, and Lowell Caneday. "Web-based GIS in Tourism Information Search: Perceptions, Tasks, and Trip Attributes." *Elsevier* (2011): 1435-437. Print.
- Duzgun, H., M. Yucemen, H. Kalaycioglu, K. Celik, S. Kemek, K. Ertugay, and A. Deniz. "An Integrated Earthquake Vulnerability Assessment Framework for Urban Areas." *Springer* (2011): 917-47. Print.
- "Election Results 2008." *Elections.nytimes.com*. The New York Times, 9 Dec. 2008. Web. 12 Dec. 2011. <<http://elections.nytimes.com/2008/results/president/map.html>>.
- "Geographic Information System." *Wikipedia, the Free Encyclopedia*. 12 Dec. 2011. Web. 12 Dec. 2011. <http://en.wikipedia.org/wiki/Geographic_information_system>.
- Gimpel, J., J. Dyck, and D. Shaw. "Location, Knowledge and Time Pressures in the Spatial Structure of Convenience Voting." *Electoral Studies* 25.1 (2006): 35-58. Print.
- "GIS Applications." *Wikipedia, the Free Encyclopedia*. 19 Oct. 2011. Web. 12 Dec. 2011. <http://en.wikipedia.org/wiki/GIS_applications>.
- Hossain, Jami, Vinay Sinha, and V.V.N. Kishore. "A GIS Based Assessment of Potential for Windfarms in India." *Elsevier* (2011): 3257-267. Print.
- Jankowski, Piotr, and Timothy L. Nyerges. *Geographic Information Systems for Group Decision Making: Towards a Participatory, Geographic Information Science*. New York: Taylor & Francis, 2001. Print.
- Jordan. "Philip Lampi - The Itinerant Historian." *H2G2 - The Hitchhiker's Guide to the Galaxy*. 7 Dec. 2009. Web. 12 Dec. 2011. <<http://h2g2.com/dna/h2g2/A55852905>>.
- Kim, J., E. Elliott, and D. Wang. "A Spatial Analysis of County-level Outcomes in US Presidential Elections: 1988-2000." *Electoral Studies* 22.4 (2003): 741-61. Print.
- Martis, Kenneth C. *The Historical Atlas of United States Congressional Districts, 1789-1983*. Ed. Ruth Anderson. Rowles. New York: Free, 1982. Print.
- Peng, Zhong-Ren. "Internet GIS for Public Participation." *Environment and Planning B: Planning and Design* 28.6 (2001): 889-905. Print.
- "Project Overview." *A New Nation Votes*. American Antiquarian Society and Tufts University. Web. 12 Dec. 2011. <http://elections.lib.tufts.edu/aas_portal/about.html>.
- Suárez, Manuel, and Irina Alberro. "Analyzing Partisanship in Central Mexico: A Geographical Approach." *Electoral Studies* 30.1 (2011): 136-47. Print.
- Tucci, Michele, and Alberto Giordano. "Positional Accuracy, Positional Uncertainty, and Feature Change Detection in Historical Maps: Results of an Experiment." *Computers, Environment and Urban Systems* 35 (2011): 452-63. Print.
- Ural, Serkan, Ejaz Hussain, and Jie Shan. "Building Population Mapping with Aerial Imagery and GIS Data." *International Journal of Applied Earth Observation and Geoinformation* 13.6 (2011): 841-52. Print.

Wu, Jilei, Gong Chen, Yilan Liao, Xinming Song, Lijun Pei, Jinfeng Wang, and Xiaoying Zheng. "Arsenic Levels in the Soil and Risk of Birth Defects: A Population-Based Case-Control Study Using GIS Technology." *Journal of Environmental Health* (2011): 20-25. Print.

Zhong, T., R. Young, M. Lowry, and G. Rutherford. "A Model for Public Involvement in Transportation Improvement Programming Using Participatory Geographic Information Systems." *Computers, Environment and Urban Systems* 32 (2008): 123-33. Print.