ETHICS ACROSS THE CURRICULUM AND GEOGRAPHIC INFORMATION SYSTEMS

Ashraf Ghaly
Union College

This module addresses spatial presentation of data using Geographic Information Systems (GIS). GIS is a powerful tool that can be used to compile large amount of data into an easy to read map that is visually appealing. Data presentation and map plotting can vary significantly depending on the method used in segmenting the given information. The wide variation of methods used in data plotting and presentation may invite manipulation that could project the information in some light that is less-than-honest to the spirit of the original data. There are many ways in which one can use the same data to plot different maps. In this module, the students are shown many examples of GIS-generated maps, with the same set of data, where each of which lead the reader to different impressions and conclusions. It is argued that map plotting should be done such that the final product offers a clear and honest presentation of the data. Any manipulation that intends to mislead the reader may be considered an abuse of the data and a violation of the basic principles of ethics. Some may argue that any method of data presentation is acceptable as long as the actual data is not altered. In response to such an argument, it has to be made clear that although the data itself is not altered, the mere fact that it is presented in a way that serves a specific purpose or diverts attention from important trends or patterns, it becomes apparent that the method of presentation's goal is unethical. Ethical presentation of data should not ignore, or avoid showing, what the data is meant to show. The following is an example of using various methods of presentation to project the same data set in various lights.

All the maps below show the results of the 2004 presidential election plotted for all the counties in the United States. This election was selected because it is the most recent one which had a close outcome. Granted, the electoral college used to elect the president of the United

Sates does not rely on the votes of counties, rather on the votes of states, the decision was made to plot these maps using county votes to sharpen the contrast between the results as the units used in plotting a map (counties) are significantly smaller in area than those of states. Therefore, instead of using fifty polygons for fifty states, every one of the maps below displays 3111 counties. Although the data for the Alaska and Hawaii counties is available, to maintain a reasonable and an easy-to-read-scale, only the results of the mainland are plotted.

Figure 1 shows the percent of votes for Bush versus those for Kerry using the natural breaks method. This map is probably the most ethical one can plot since the map maker has, arguably, little room to misuse the data. This method does not require human interference to plot it. When this method is selected, the only input required is the number of classes, which is five in the case of the map plotted in Figure 1. The computer utilizes this answer to analyze the data and divide it statistically into the specified five classes. A red area of the map indicates a republican stronghold, whereas a blue color indicates dominance of democrats. Other colors show counties with split vote.

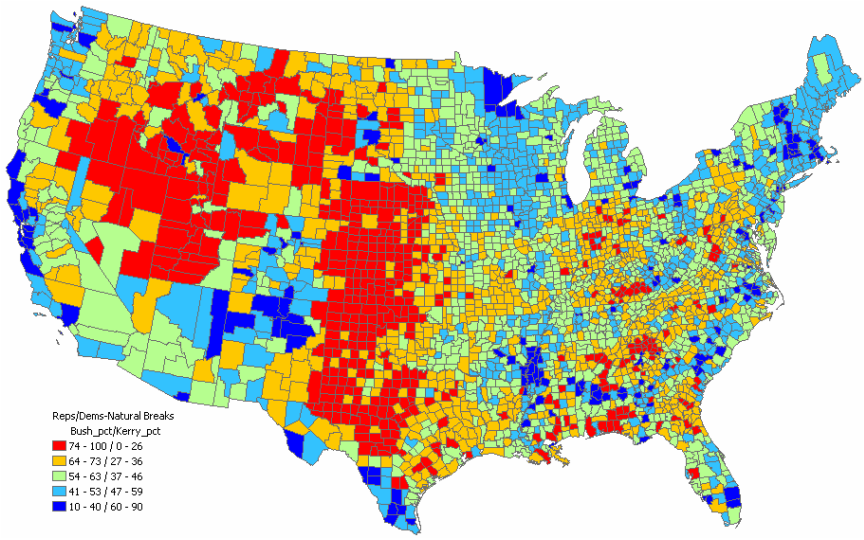

Figure 1. Percent of votes for Bush/Kerry using the natural breaks method.

Figure 2 shows the percent of votes for Bush versus those for Kerry using the manual method. In plotting this map, the user can manually specify the number of classes and boundary of each. For the sake of comparison, the number of classes is maintained constant at five for all the maps plotted in this example. Each class is given an equal 20% weight. Comparing this map against the one shown in Figure 1 reveals almost total disappearance of democratic areas as represented by the blue color and the considerable shrinkage of republican areas signified by the red color. This map can be significantly manipulated by manually specifying the limits of each of the five classes. It can be reproduced to make one of the colors most dominant, which can give a false impression of the election results.

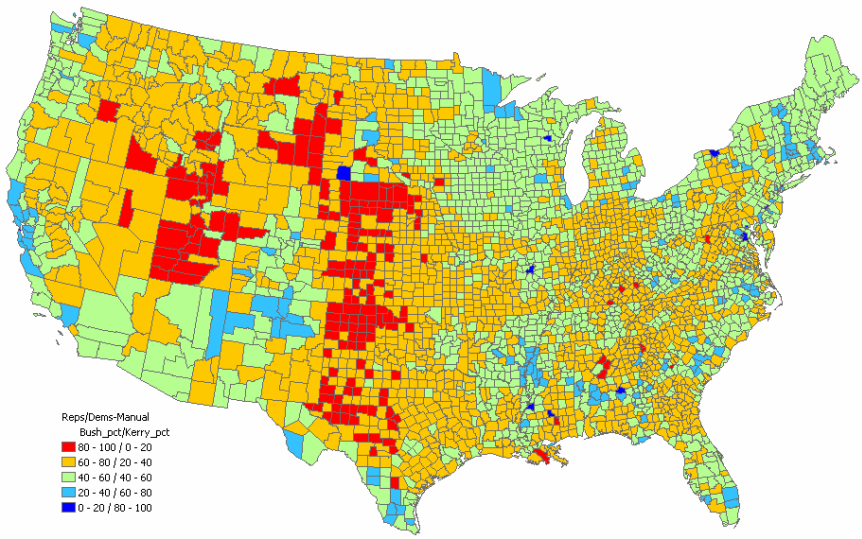

Figure 2. Percent of votes for Bush/Kerry using the manual method.

Figure 3 shows the percent of votes for Bush versus those for Kerry using the equal intervals method. In this method, the mapmaker is required to specify the number of classes, which is maintained at five similar to other maps. Rather than dividing a total of 100% into five classes, the equal intervals method uses the widest range of data, which is 90% in this case, and divides it into five classes; 18% each. The resulting map is very much similar to the one shown in Figure 2 but is appreciably different from the one shown in Figure 1. Manipulating this type of map can easily give the reader erroneous conclusions. This can be done by increasing or decreasing the number or classes or by assigning colors that will project one party or another in a favorable light.

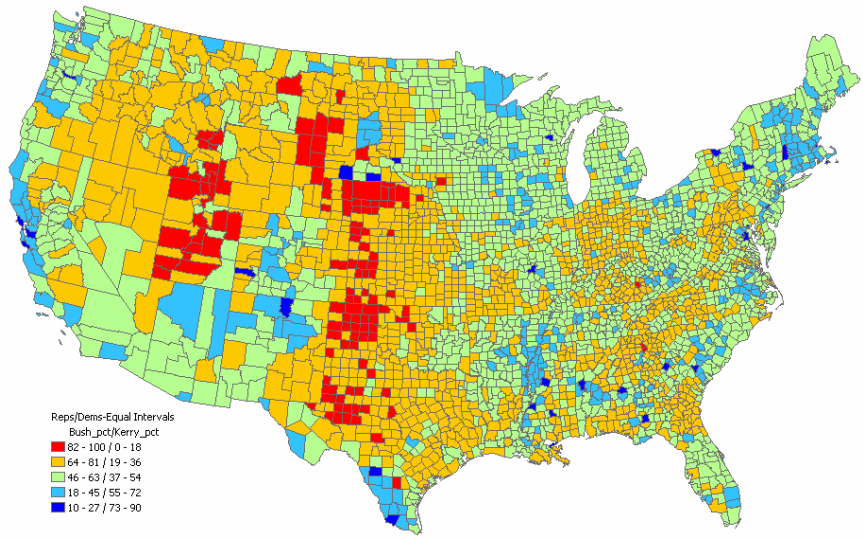

Figure 3. Percent of votes for Bush/Kerry using the equal intervals method.

Figure 4 shows the percent of votes for Bush versus those for Kerry using the geometrical interval method. Although the data set used in plotting this map is the same used in previous ones, it is evident that this method gives the impression that Bush was dominant in many counties as the red color is prominent in the map. One must also take into consideration the fact that the area of the counties in the Western and Midwestern parts of the United States are significantly larger than those in the Eastern part of the country but also appreciably lighter in population density. Since people's votes, not areas, determine election outcome, another factor that must be accounted for is population density. This, however, is out of the scope of this module, but it should be mentioned here that all the maps shown in this module would be subjected to significant distortion if the areas of counties and states were sized to account for population density.

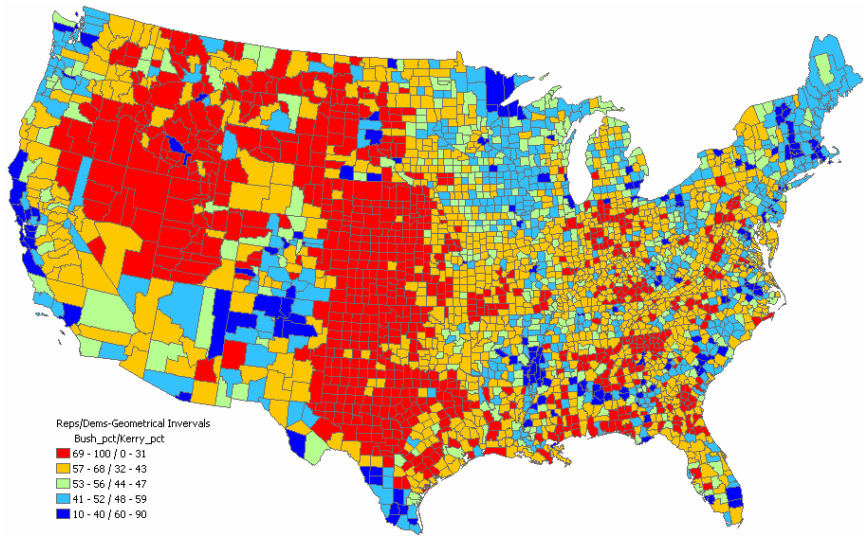

Figure 4. Percent of votes for Bush/Kerry using the geometrical intervals method.

Figure 5 shows the percent of votes for Bush versus those for Kerry using the quantile method. The data in this method is divided to five classes as was done in previous maps to maintain consistency. This method produces the sharpest contrast between republican and democratic regions. The reason is, by its nature, this method of map making concentrates on areas where votes for one party or another are heavily clustered. It does not place as much emphasis on areas where votes, for one party or another, are light.

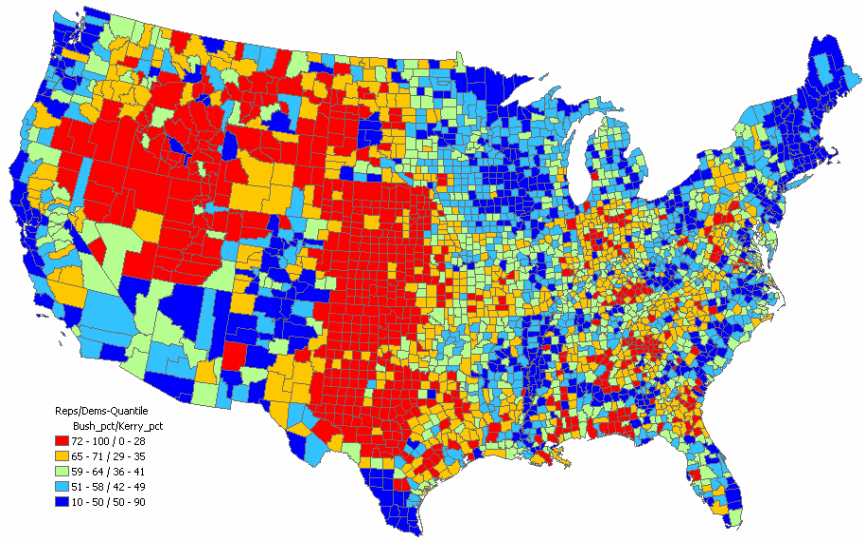

Figure 5. Percent of votes for Bush/Kerry using the quantile method.

Judging from the distribution of colors shown in each of the above maps, one gets different impressions and draws different conclusions. As stated earlier, all these maps were produced with the same exact set of data but with different methods of map making. Depending on the goal of the mapmaker, it is shown that producing maps to serve a specific purpose is possible. The argument this module attempts to make is: it is unethical to mislead the reader by using a specific plotting method with the intention to project a certain factor or a given phenomenon in a favorable or unfavorable light. Data should be presented for what it is and should be plotted in a way that reflects its spirit.