Visualizing Lynching Data of Ten Southern States: A Prototype

Cynthia Lovett

Information Visualization [17:610:525:85]

Rutgers University

School of Communication and Information

Dec 20, 2011

**Introduction**

“Visualizing Lynching Data of Ten Southern States” is an information visualization prototype created to discover ways of representing the Lynching Era of the American South at the turn of the 20th century. Stewart E. Tolnay, Professor of Sociology at the University of Washington, generously provided the dataset for this project. Tolnay, along with colleagues, has written numerous seminal works in the study of mob violence, demography, and the Great Migration.

This project emerged from an earlier geographical visualization assignment using the same dataset, in attempt to explore the categorical aspects of this historical data.

**Historical Context of the Lynching Era**

Lynching is defined as “extrajudicial execution carried out by a mob, often by hanging, but also by burning at the stake or shooting, in order to punish an alleged transgressor” (Wikipedia, <http://en.wikipedia.org/wiki/Lynching_in_the_United_States>).

Another definition is “to put to death, especially by hanging, by mob action and without legal authority” (<http://dictionary.reference.com/browse/lynching>).

The “Lynching Era” refers to the late 19th and early 20th century in the United States, when conditions following the Civil War and the abolishment of slavery, produced an atmosphere of social and economic tension that may have contributed to thousands of lynchings. Many of these incidents occurred in the 1890s, which has been described as “the bloodiest decade” (Baily, Tolnay, Beck, Roberts & Wong, 2008).

The specific causes of mob violence during this period vary as discussed by Baily et al. (2008):

Although complete consensus regarding the most important antecedents of lynching may not yet exist, previous comparative studies have indicated that southern lynchings (primarily of African Americans) were more likely to occur when (1) the black population was proportionately larger, (2) the white population suffered economic hardship, (3) the Democratic Party was stronger, (4) the white population was threatened economically or socially by the black population, (5) the level of black outmigration was lower, and (6) alternative legal sanctions for serious crimes were lacking (p. 48).

While the numbers victims of lynchings during this period were disproportionally African-American, whites, immigrants, and “outsiders”, who were either accused of crimes or whose actions threatened the social conventions that were shifting dramatically, were also targets. Brundage (1993), explains why white lynchings occurred during between 1880 and 1940 in Virginia and Georgia:

The lynchings of whites in the two states represented the perpetuation of at least a century-old tradition of punishing whites who deviated from community standards of behavior, whether by abusing their spouses and children or by holding unorthodox moral, social, or political beliefs (Brundage, 1993, p. 87).

**Overview of the Beck & Tolnay (2004) dataset**

The Beck & Tolnay (2004) dataset is built on a number of sources including the annual list of lynch victims from the *Chicago Tribune*.  It contains confirmed data on 2805 individuals: including the name, date, location of lynching (state and county), sex, race of victim, race of mob, and the accusation or offense. The ten states the dataset covers are:  Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, and Tennessee.

For this visualization project, this dataset was selected because of the granularity of victim details. It was provided in the form of a Microsoft Excel spreadsheet with the eight categories of data mentioned.

Lisa D. Cook, Assistant Professor of Economics at Michigan State University, discusses numerous lynching datasets in her article, “Converging to a National Lynching Database: Recent Developments.” Cook (2011) explains that the work of Tolnay and Beck grew from careful examination of data from the Tribune, NAACP, and Tuskegee University. “The Tolnay and Beck (1995) data have been extended to include more characteristics of the victims of the community to examine common features of lynching victims, their households and communities” (p.11). Although the Beck & Tolnay (2004) dataset not contain community and household data, it would be useful to examine these factors in a future project.

Even with the meticulous confirmation of details in the Beck & Tolnay (2004) dataset, there are several victims for which there is no name provided that may be described in vague terms such as “Unknown”, “Unknown Chinese”, or “Unknown Negro.” The race of each victim was recorded as “black,” “white,” “unknown,” or “other”; the sex was recorded as “male,” “female,” or “unknown.”



Figure 1. Beck & Tolnay (2004) dataset in Microsoft Excel format.

Confirmation of the racial identity of victims is difficult due to the rigid categorizations of racial groups during this time period. Non-white and non-black victims could be classified as either black or white. In some cases, race, sex, and location (county level) were undetermined, but in the majority of cases, race is classified as either black or white.  Cook (2011) discusses how the caste system of the postbelllum and Jim Crow eras created challenges for determining the race of individuals: “Each non-black victim, i.e., white, Chinese, Hispanic, Italian, or Native American, was recorded as white” (p.10-11).  Bailey, Tolnay, Beck, Roberts & Wong (2008) explain that black victims were often recorded as “Negro”, “Colored”, or “Mulatto” (p. 52).  Therefore, confirming the racial identity of victims required additional research.

**Representing Data: Geographic, Categorical & Nominal Types**

The present visualization was in part, inspired by an earlier project that served as an initial exploration of the Beck & Tolnay (2004) dataset. Using Google Fusion Tables, on online-based tool for visualizing data, the state and counties were geocoded in order to prepare the data for map display. The original dataset was reorganized into separate files for each of the ten states in order to determine the number of victims per state and county according to race and sex.  Afterwards, it was merged the Google (*KML)* file of [United States Counties](http://www.google.com/fusiontables/DataSource?dsrcid=210217) which contains geocoded data.  Because the dataset had such a small number of victims categorized as “Other,” they were combined with the" unknown" category in a new group called  “Other/Unknown.”  Incidents where the county was “undetermined" were not included.

The basic functions of the Google Fusion Tables software are equipped for the representation of quantitative data and geographic locations. However other kinds of nominal data, such as names, offense, and race were difficult to translate into Google Fusion Tables. In order to represent these aspects, the info-window was programmed dynamically to include Google Charts for other categories.

The info-window pulled quantities from a newly created column in the spreadsheet representing total counts by race and sex in each county. These quantitative values were calculated separately because the original dataset was organized by victim name with text values for each category (i.e. “Blk”, “Wht”, “Unk” or “Male”, “Fem”, “Unk”).



Figure 2. Modified Beck & Tolnay (2004) dataset with quantitative values.



Figure 3. Google Fusion Table Project of Tolnay & Beck Lynching Data with Info-window.

Many observations could be made from the Google Fusion Table maps visualization. Analysis of the dataset and the geographic visualization produced the following findings:

* Out of 2805 lynchings recorded, 2462 were black and 287 where white.  Four were classified as other, and 42 were unknown or undetermined.
* In terms of sex, 2701 were male and 79 were female. There were 25 individuals for which the sex was unknown or undetermined.
* The states with the highest incidents of white victims were Louisiana (53), Arkansas (48), and Kentucky (37).
* The states with the largest number of female victims were Mississippi (18) and Georgia (13).
* The two states out of the ten with the highest number of lynchings were Mississippi (538) and Georgia (438).
* The smallest number lynchings were evident in North Carolina (97) and South Carolina (151).

Google Fusion Tables allowed for flexibility in color-coding and so a few alternate visualizations were created to highlight lynchings at the state and county level.



Figure 3. Color coding comparison at the state and county levels in Google Fusion Tables

The creation of the Google Fusion visualization was an informative process for exploring the Beck & Tolnay (2004) dataset. However, one of the unique and interesting aspects of this data are the very specific details on individuals, such as the names and offenses. It was clear that to attempt to represent these nominal aspects of this dataset, other tools would have to be explored.

**Tool Selection & Overview of Flash Builder 4.5 Flex Framework**

After reviewing several visualization tools such as Google Charts & Google Fusion, IBM ManyEyes, and Tableau software, it was decided to develop the prototype using ActionScript 3.0. Two choices available for programming environments were Flash Professional and Flash Builder 4.5. It was decided to work with Flash Builder because the accompanying Flex framework that runs with mxml and ActionScript 3.0 contains several charting and visualization tools. In addition, at the time of this writing, Adobe offers a free version for students and faculty on the Adobe Developer Connection website ( <http://www.adobe.com/devnet-apps/flex/free/> ).

The “Tour de Flex” application from Adobe (<http://www.adobe.com/devnet/flex/tourdeflex.html>) provides an overview of the kinds of visualizations that can be created with Flex through Flash Builder, therefore having these options available during the development process would be advantageous.

There are a number of Open Source applications available online that require coding experience with various languages. Flare Visualization for the Web (http://flare.prefuse.org/), developed by of UC Berkeley Visualization Lab, contains ActionScript libraries for developing online visualizations. After downloading the most recent source files (2008), it was found to be incompatible with the recent version of Flash Builder 4.5 and for that reason it was not used for this project.



Figure 4. Adobe Tour de Flex highlights Flex capabilities with sample projects.

After this exploration of different tools, it was decided to hardcode the visualization in Flash Builder as a means to increase familiarity with ActionScript 3.0 and to have the option of customizing the visualization from scratch.

**The Design & Programming Process**

To begin the process of programming and designing the visualization, many basic ActionScript projects were used as models to assist in learning the programming language. At the same time, sketches were produced for brainstorming the possibilities. In some of the early sketches the states were structured as a bar composed of numerous squares, which represented individuals.

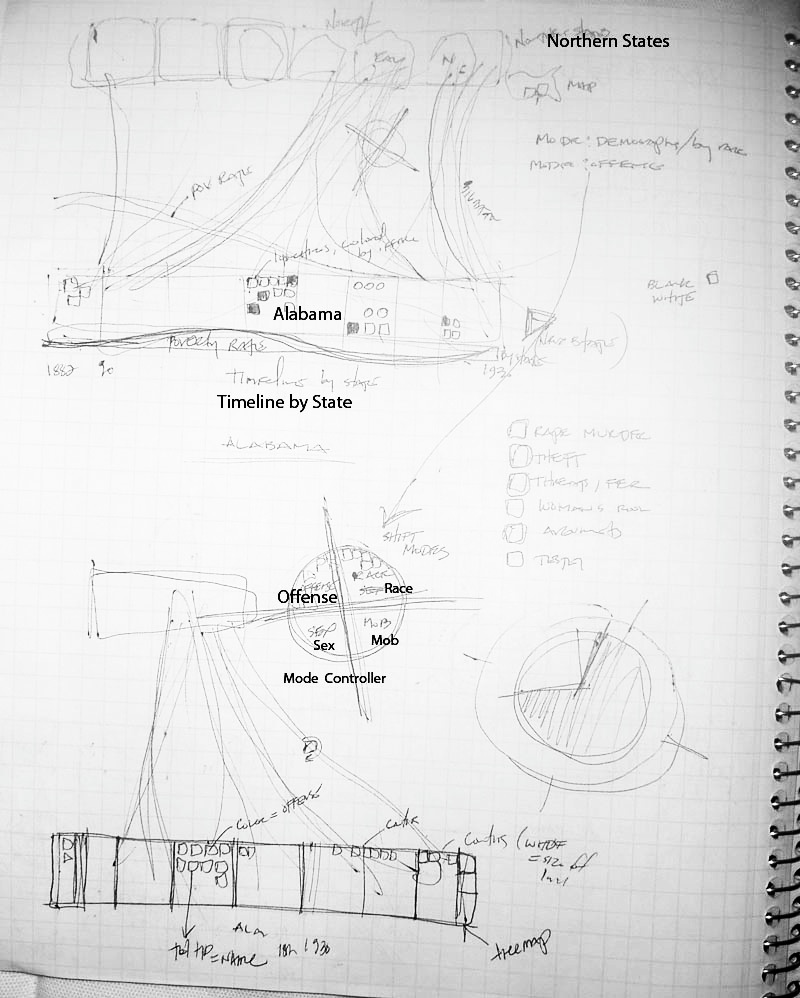


Figure 5. Concept sketch for visualization of lynching, migration, and poverty rate

The initial concept also included migration and poverty data, however the prototype at this stage, does not include these aspects.

The next step was to determine how to import ten Microsoft Excel spreadsheets into Flash Builder. Because Flash Builder works with MXML for its display and interface language, it followed that the spreadsheets should be converted to XML format for greater compatibility. Also, at this time, Flash does not work well with Excel files. Another option was to enter the data as an associative array object in ActionScript, but with thousands of entries, this could quickly become unmanageable. Having the data compact and well contained in a separate XML file seemed to be the best option. Although Excel allows files to be saved in XML format it also produced many lines of code and namespaces that appeared to be proprietary to Microsoft. In the end, an online program called Mr. Data Converter (<http://shancarter.com/data_converter/>) was used, which has the capability of copy-and-pasting directly from the spreadsheet as well as the ability convert it to two possible formats: Node and Properties. The Nodes option was the chosen layout for this project.

To import XML files into ActionScript the URLLoader Class was used. The URLLoader (<http://help.adobe.com/en_US/FlashPlatform/reference/actionscript/3/flash/net/URLLoader.html>) downloads text and XML files for use in data-heavy applications. In addition, ActionScript 3.0 has an XML Class available with numerous methods for parsing data and translating them to Strings if needed.

****

Figure 6. Flex FB interface with MultiXMLLoader Class. includes a side panel with project file structure.

The programming and design of the visualization developed alongside the learning process. The main class file is called MXML\_MultiXMLLoader, which is the MXML file that serves as a wrapper for the project. This Class calls the MultiXMLLoader class, which loads the ten XML files—one for each of the ten southern states. After loading the files they are parsed and built into a multidimensional array, which includes screen locations (x, y). Next, the MXML\_MultiXMLLoader calls the VictimSquare class and passes the data through parameters so it can create the display objects and add EventListeners to register MouseEvents. The MouseEvents trigger changes in the display objects (the VictimSquare Objects) and parse the data for specific properties. There is also a class called BackgroundImage which is called from the XMLLoader Class to draw the grid in the background and label each of the states. The GridLegendText Object formats any text on the screen, and the GridLegend Object controls the display of the Legends with color-coding properties.

The main class, MXML\_MultiXMLLoader was used mainly for key interface controls such as the panel, containers, and scrollbars or what is termed “skins”. Some ActionScript code was written inside the MXML file to send events back and forth between the MXML main class and the ActionScript classes, but the bulk of the visual aspects were drawn with ActionScript. One of the key discoveries was that when working with MXML container and panel objects along side ActionScript (AS) classes, the AS files need to “extend UIComponent” instead of “extend Sprite” in order to appear on the screen.

**The Visualization Prototype & Observations**

The prototype created represents a first attempt of programming in ActionScript 3.0 and emerged organically during the process of programming. The horizontal axis displays the ten states along a row alphabetically, similar to a traditional vertical bar chart. The bars are broken up into 10 x 10 pixel squares with each one representing an individual lynching victim. Along the vertical direction there is a grid that displays numbers at intervals of 100 that rises up to 500, with Mississippi having highest number of lynchings (538). The opening screen displays all the 2805 victims in a uniform gray color and provides brief instructions for users on how to explore data using the drop down menu.



Figure 7. Introduction states number of total victims, displaying squares uniformly.



Figure 8. Dropdown menu displaying 5 filtering options

The dropdown menu contains five ways to filter or view the data: “Race of Victim”, “Race of Mob”, “Sex of Victim”, “Offense”, and “Timeline.” When users select an option the squares are color-coded with different schemes to represent the categorical data.

Users can also use the mouse to roll over the individual squares to produce a pop-up tool tip and discover details such as the exact date of the lynching, the county and state, as well as the specific offense.

The “Race of Victim” screen displays black victims in black, white victims in light gray, and unknown victims in orange. As discussed earlier, some of the racial classifications are speculative because of the way race was conceived during that time period. As expected, the race of the victims is overwhelmingly black, with some white and unknown victims.



Figure 9. "Race of Victims" shows more white lynchings in Arkansas and Kentucky.

From the “Race of Victim” view, it can be noted that most of the lynchings of white and race-unknown victims occurred closer to the turn of the century because they are at the bottom of the screen (the squares were arranged onscreen with the earliest dates at the bottom moving upwards to 1930 at the top of each bar). It can be observed that more lynchings of white and unknown victims occurred in Arkansas and Kentucky.

The “Race of Mob” screen shows an almost reverse image of the “Race of Victim” screen, although there is a sprinkling of black mobs that lynched both black and white victims. There were no “unknown” or “other” mobs listed in this dataset.



Figure 10. "Race of Mob" shows mostly white mobs and very few black mobs. No unknown mobs



Figure 11. "Sex of Victim" shows few female victims in blue with some child victims among the unknown.

The “Sex of Victim” screen displays a similar configuration to the first two sections. The victims were nearly all male with some female victims. One could also discover that there were child victims among the unknown-sex category.

The “Offense” screen displays the alleged offenses in five different categories: “Murder” or “Assault,” “Rape,” “Robbery,” “Arson,” and “Miscellaneous.” Out of 2805 lynchings in the dataset, there were about 270 different offenses or accusations. Programming was used to parse the offense property to find words matching the category. But even with programming, a lot of human thought was needed to make some decisions on where many offenses fell.



Figure 12. "Offense" screen displays a confetti of 270 reasons filtered into 5 main categories.

For those offenses that were twofold, such as “Train wrecking and Murder”, or “Robbery and Murder”, it was decided to classify them by the more serious offense. Within the five categories, different hues or saturation levels of that same color represented the intensity of the accusation. For example a victim lynched for arson would be displayed in a deeper yellow, whereas a victim lynched for being the “Mother of Arsonists” was displayed in a very light or pastel yellow. Arson was a category that included all kinds of property damage such as “Injuring Livestock”, “Poisoning Mules”, “Barn-burning”, “Train wrecking”, or “Cutting Levee.”

The light blue color represents those offenses that were not easily classified into the other four categories, some of them include: “Entering a girls room”, “Arresting white cappers”, “Refusing to aid lynch posse”, “Angered Klan”, “Frightened woman”, and “Testifying.

Looking at the “Offense” view, there is confetti of reasons depicted as a scattering of colors with all kinds of offenses mixed in. It appears that a large number of victims lynched for “Attempted Rape” displayed in the bright orange color. There may also have been a significant number of murder or “Attempted Assault” related victims. The “Arson” category was less significant followed by “Robbery” accusations. The Miscellaneous category figures predominately in the later years in both Kentucky and North Carolina where there are large patches of blue at the top of the bars.

The timeline screen displays decades of between 1882 and 1930 in different colors. Since the VictimSquare objects were loaded with earlier years at the bottom and later years at the top, jagged bands of color provide a snapshot of the quantities per decade.



Figure 13. "Timeline" view shows the largest numbers of lynchings in most states occurring 1890 -1900.

The timeline confirms that the largest numbers of lynchings occurred in the last two decades before the turn of the twentieth century with numbers particularly high in Louisiana and Mississippi between 1890 and 1900. North Carolina shows numbers that were much lower than the other states overall and especially during this decade. In the following decade (1900-1910) all states experienced a slight decline in lynching numbers except for Mississippi with numbers between 1900-1910 being almost equal to the prior decade.

**Future Implementations & User Feedback**

This visualization prototype represents a starting point for learning ActionScript 3.0, and an exploration of ways to apply color-coding to categorical data, as well as method for discovering how to represent historical data. The initial concept included migration and economic data and a future version of this project that represents these aspects might broaden the context of this study show the larger effects of the Lynching Era on the social landscape of the United States.

A few users have tested the project briefly and provided some basic feedback. One user looked at both the geographic visualization as well and said, “Both graphs(?) are excellent and relay the same information but in a slightly different way.  But they are both comprehensive and easy to use.”

Another person, an art professor said: “Your project looks good and basic information is very easy to access/ find.   The only thing I found myself wishing for is more "exciting" visual representation of the info.   This stems from all the exciting data visualization I've seen lately.  This desire may not be appropriate for your work.”

A more experienced programmer of both games and analytics said:

"Timeline" is definitely a missed opportunity:  By looking at the graphs and "figuring out" the relative sizes, I can see that TN noticeably improved in the 1890, 1900, 1910 decades, and GA did not. But, the fact that I have to "work at it" is something of a failure point.  Maybe if the decades were lined up together, it would be easier to see. Perhaps:  states are still horizontal across the bottom, decades are vertical bands up the y-axis.

After learning the fundamentals of the ActionScript language it is now easier to see how the visualization could be expanded and re-designed. It would be interesting to explore different configurations of the squares, or to animate their position depending on user selections. The “Offense” section could benefit from another filter that allows users to see just one “Offense” category at the time, while disabling and graying-out the other squares so the numbers could be viewed more clearly. The search function, which would allow users to find names of potential family members, is another feature that was not fully programmed in this build.

Other ideas include the use of the search function for highlighting aspects such as specific offenses. This way if a user entered train wrecking,” all related squares would be highlighted or increase in size.

Finally because the squares animate and expand on the rollover of the mouse, the “hit area” also increases size and users cannot easily access neighboring squares because they are so small. Decreasing the “hit area”, or else employing a fish-eye effect of increasing the size of neighboring squares might allow users to trigger them without difficulty.

**References**

Bailey, A. K., Tolnay, S. E., Beck, E. M., Roberts, A. R., & Wong, N. H., (2008). Personalizing lynch victims: A new database to support the study of mob violence. *Historical Methods, 41*, 47-61.

Beck, E. M. and Stewart E. Tolnay. (2004). Confirmed Inventory of Southern Lynch Victims, 1882-1930 [Data File]. Machine-readable data file available from creators.

Brundage, W. F. (1993). *Lynching in the New South: Georgia and Virginia, 1880-1930*. Urbana: University of Illinois Press.

Cook, L. D. (2011). Converging to a national lynching database: Recent developments. Under review, *Historical Methods*.  Retrieved from: <https://www.msu.edu/~lisacook/hist_meths_lynch_paper_final.pdf>

**Resources**

Center for the Studies in Demography And Ecology. Web page of Dr. Stewart E. Tolnay.

<http://csde.washington.edu/people/interests.php?id=47>

Website of Professor Lisa D. Cook, Assistant Professor of Economics and International Relations Michigan State University

<https://www.msu.edu/~lisacook/>

Wikipedia Lynching Article

<http://en.wikipedia.org/wiki/Lynching_in_the_United_States>

**Programming and Development Resources**

Adobe Developer Connection (Free Flash Builder for students and faculty)

<http://www.adobe.com/devnet-apps/flex/free/>

Flare Data Visualization for the Web

<http://flare.prefuse.org/>

Tour de Flex, Adobe

<http://www.adobe.com/devnet/flex/tourdeflex.html>

Mr. Data Converter

<http://shancarter.com/data_converter/>

Adobe 3.0 Actionscript Reference: URLLoader

<http://help.adobe.com/en_US/FlashPlatform/reference/actionscript/3/flash/net/URLLoader.html>

StackOverflow, Programming Forum

<http://stackoverflow.com/>

Actionscript.org, Programming Forum

<http://actionscript.org/>