Sam Behrens

SDS 235

December 16, 2015

Visual Analytics Reflection

The project that I chose to didn’t seem to be inherently challenging because I knew that there was distinct connections between the blues and the great migration. I also knew the technology to create this visualization like timelines and maps existed and could easily be created. However, what I did not expect or recognize is that historic data doesn’t come in easily accessible formats and there had to be a lot of data processing to get the data in a format that I could easily read.

This data processing includes scraping at least 6 different websites for timeline data and I had to scrap them differently because the timeline was formatted in different ways. Sometimes it was in a table with those associated tags, sometimes it was a list, sometime it was paragraphs with dashes, and so on. Therefore I had to write several python scripts to deal with the many ways that one could create a timeline and then convert that into a json file. Furthermore, finding historic census data seemed to be easy but it turns out that the census API does not contain historic data from the 1900’s. Also how the U.S government thought about race and what population data was thought to be important changed drastically from 1900 to 1970 so that I could not find consistent formatting. Since this proved to be challenging I had to find data from which I could calculate African American populations from the breakdowns of race, age, and sex. All this made the estimated project time longer than expected and I did not have time to do the work to create the logical connection between the great migration and the blues such as including cities, state lines, event locations, bibliographical artist information, discography, and song samples.

Despite all of the roadblocks in this project I found it to be pretty interesting and enjoyable. I learned so much about data processing, data sets, and most surprisingly I learned so much about American history just through looking at how the data was saved. What I found to be most curious was the census. What the census thought about race throughout the 19th century changed drastically. Often early data sets would contain information on white, black, and other with no distinctions of other races until the 1950’s and 1960’s. It was also interesting to see that the census took note of white native born, white foreign born, and then just black. This seems like the census assumes that all black people were born in the U.S or it just doesn’t care that there are foreign-born black people within the U.S. Also the phrasing around race is interesting as the terms negro, colored, and black appear at distinctly different times.

I think you cannot draw conclusion around the blues and the great migration from this visualization but I think the timeline is interesting and I enjoy the way that the map works with the timeline. I think I will continue working on this project to get it to a point where you can really see those connections and be able to explore the blues through this interface. I think it could be really valuable and interesting.