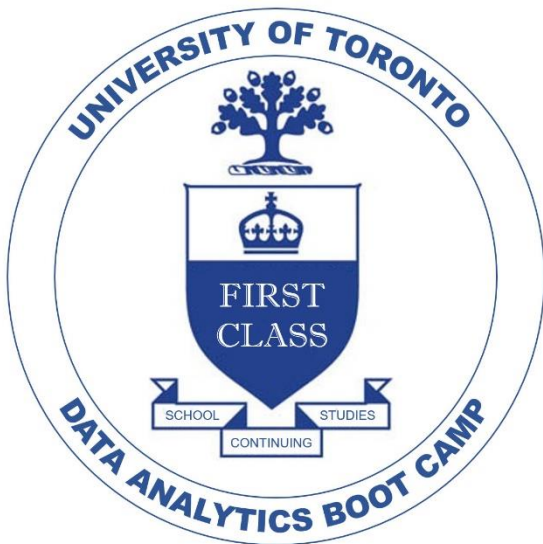


Death American Style



2019 MAY 01

University of Toronto SCS
Data Analytics Boot Camp Final Project
Jose Tomines



Death American Style

Death American Style is an interactive visualization correlating Gun Law Grade for each state of the US with firearm deaths, suicides, mass shooting incidences, hate groups and hate crimes. Submitted as the Final Project for the University of Toronto School of Continuing Studies Data Analytics Boot Camp, this project employed several different techniques in grabbing and manipulating data from several different sources, in order to graph and to provide web-based visualization.

“The greatest value of a picture is when it forces us to notice what we never expected to see.” - John Tukey

Though **Death American Style** was inspired to provide meaning and insight to very polarizing and American issue of gun rights and gun control legislation, this document will not touch on the actual analysis of the data and any conclusions provided. This document serves to record the process used to assemble all the components of this project. This also provides my thoughts and decisions as I was bringing this website to fruition.

Note that **Death American Style** is a project that I have every intention of improving and expanding. I will attempt new ways to visualize the data, including employing different tools and visualization modules. Because of these lofty intentions, this document will also be a living and breathing document.

Methodology

The original proposal (see Reference Documents section) suggested the following steps:

- 1) Determine specific data to be used
- 2) Determine visualization of data
- 3) Determine framework of website
- 4) Build data collection and cleanup code
- 5) Build DB
- 6) Build visualization code
- 7) Analyze and comment on visualization
- 8) Assemble website

The final project for the boot camp was meant to demonstrate a student's overall knowledge through the process of creating a product that was meaningful to the student. Though there were no requirements for implementing this project, the original set of steps proposed was based on providing the ideal data visualization website.

The overall schedule and duration to complete this final project was difficult because of assignment work needed to "graduate" from this boot camp, and other conflicting priorities. Because I needed to work on my own schedule, I decided to deliver this project by myself instead of working in a group as I had done previously. These shifting priorities also led to decisions of not fully providing the ideal mechanisms for some of these steps that was previously intended in order to meet the project delivery deadline.

Regardless of the changes from the ideal plan, this project was delivered in time to successfully "graduate" from this boot camp with the deliverable earning an A+ grade.

Determining Data And Visualization

During the project proposal stage, I needed to provide a visualization proposal based on ample available data. My previous team projects always centered on current event issues, with the understanding that though there is ample sources of different data available publically, not all data matches your needs cleanly. But I had discovered some very powerful websites centered around 3 areas with ample data: 1) death statistics by guns including suicides and mass shootings, 2) state gun laws and how they were evaluated, and 3) hate groups & hate crimes.

These were the sources identified:

1. Giffords Law Center - To Prevent Gun Violence
(<https://lawcenter.giffords.org/scorecard/>)
2. Inverse - 7 Maps to Explain U.S. Gun Control Laws, By State
(<https://www.inverse.com/article/37141-state-gun-law-maps>) – Project about measles and vaccine rates by user artichoke
3. KQED News – Are States with Tough Gun Laws Actually Safer
(<https://www.kqed.org/lowdown/6286/are-states-with-tough-gun-laws-actually-safer>)
4. Guardian News – Gun Laws in the US
(<https://www.theguardian.com/world/interactive/2013/jan/15/gun-laws-united-states>)
5. CDC - Firearms Mortality
(https://www.cdc.gov/nchs/pressroom/sosmap/firearm_mortality/firearm.htm)
6. FBI – Crime in the US (<https://ucr.fbi.gov/crime-in-the-u.s/2017/crime-in-the-u.s.-2017/tables/table-4/table-4.xls#overview>)
7. ABC News – Mass Shootings in the US
(<https://www.abc15.com/news/data/mass-shootings-in-the-united-states-when-where-they-have-occurred-in-2019>)
8. Southern Poverty Law Center – Hate Groups Across the US
(<https://www.splcenter.org/hate-map>)
9. <https://www.motherjones.com/politics/2012/12/mass-shootings-mother-jones-full-data/>

After these sources were identified and with much of the data being state-centric, it was decided that using choropleth maps ¹to visualize 2 different data issues per state was a strong way to present the data.

It should also be noted that in order to implement the visualization, there was a need to use data that were not directly associated to the issues being explored. This included:

10. State geocentric coordinates required to display state specific data from above (Wikipedia: https://en.wikipedia.org/wiki/List_of_geographic_centers_of_the_United_States), and
11. State population data to normalize the values of aggregated data (United States Census Bureau: <https://www.census.gov/data/datasets/time-series/demo/popest/2010s-national-total.html>)

Determining Framework of Website

Based on the schedule, I had decided to leverage the same framework that was implemented in 2 of my previous group projects. Using bootstrap implemented on static webpages, with a common navigation bar implemented on all pages. The home page would incorporate a moving carousel of provocative pictures, with each linking to a different visualization page of the issue. Each page would have its own choropleth map correlating 2 different data sets, with a slider to move through the different years of the data.

A last-minute decision was made to also add a Summary page to show graphs to visualize the different analysis correlating the issues.

¹ a map that uses differences in shading, coloring, or the placing of symbols within predefined areas to indicate the average values of a property or quantity in those areas.

Extracting and Cleaning Data

Many of the above sources included the functionality to download the data as a csv or an excel file. Though I could have used web-scraping to gather data embedded in the html, some pages, including the state geocentric coordinates, were small enough to use a simple copy and paste of the displayed table into an excel spreadsheet. For data first downloaded into excel, I used VBA to properly parse and format the data, sometimes automating the process as a macro when dealing with multiple files. The goal was to format the excel data so that it can be converted to csv files for use by Python scripts to further clean up and manipulate data.

The following lists how each data source was extracted, and cleaned up:

Building DB

The original plan was to create a SQL database and use Flask to serve up the data to JavaScript files, following what was used in Project 3. Unfortunately discussing the pitfalls experienced in the project with the project members, I realized that the schedule didn't allow me to properly create the flask engine and go through the debugging process. I decided to have the website pull the data directly from the properly formatted csv

Web Visualization

The original plan was to create a SQL database and use Flask to serve up the data to JavaScript files, following what was used in Project 3. Unfortunately discussing the pitfalls experienced in the project with the project members, I realized that the schedule didn't allow me to properly create the flask engine and go through the debugging process. I decided to have the website pull the data directly from the properly formatted csv

Analysis and Summary

The original plan was to create a SQL database and use Flask to serve up the data to JavaScript files, following what was used in Project 3. Unfortunately discussing the pitfalls experienced in the project with the project members, I realized that the schedule didn't allow me to properly create the flask engine and go through the debugging process. I decided to have the website pull the data directly from the properly formatted csv

Web Application Flow

The original plan was to create a SQL database and use Flask to serve up the data to JavaScript files, following what was used in Project 3. Unfortunately discussing the pitfalls experienced in the project with the project members, I realized that the schedule didn't allow me to properly create the flask engine and go through the debugging process. I decided to have the website pull the data directly from the properly formatted csv

Reference Documentation

The original plan was to create a SQL database and use Flask to serve up the data to JavaScript files, following what was used in Project 3. Unfortunately discussing the pitfalls experienced in the project with the project members, I realized that the schedule didn't allow me to properly create the flask engine and go through the debugging process. I decided to have the website pull the data directly from the properly formatted csv