The Visualization of Gun Violence in the US between 2013-2018 with Geo-Map, Hoover Charts and Cross-Filter

By

Xinyao Han (xh1082@nyu.edu), Zoe Xiao (yx1750@nyu.edu)

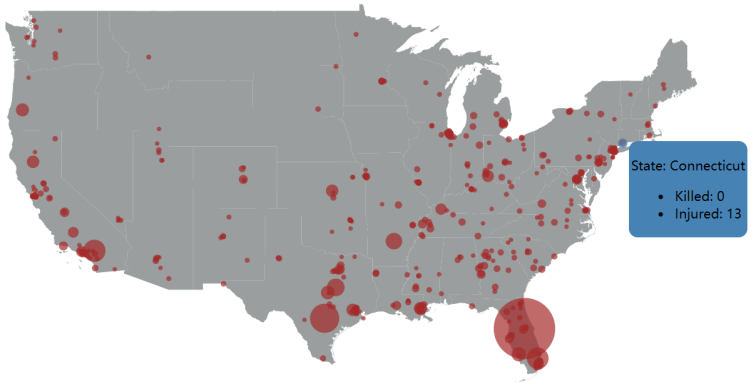
NYU Shanghai

May 2022

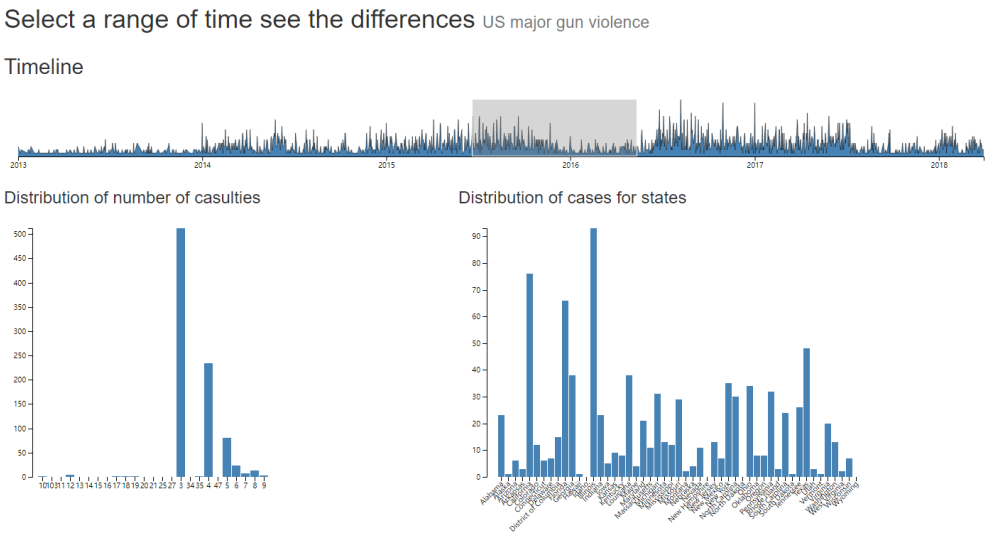
Overview

This project provides a visualization system to visualize gun violence in the U.S within the timespan of 2013 to 2018. Our purpose is to provide a vivid and customized descriptive analysis on gun violence and therefore assist those engaged in discussions and activities concerning gun violence, including analysis of proposed regulations or legislation relating to gun safety usage.

This visualization system demostrates major gun violence cases across the U.S with a geographical map and cross-filter bar charts. The geographical map visualizes geographical features of gun violence cases, proposes detailed case information with a hoover informatoin box, and redirects users to crime report webpage with on-click action.



The cross-filter bar charts provide customized analysis regarding gun violence timespan and according gun violence distribution across states and number of death and injuries.



Dataset Description and Data Processing

1. Dataset Description

We'll use a publicly available dataset on Kaggle, which is a comprehensive record of over 260k U.S gun violence incidents from 2013 to 2018. Each record has 29 attributes that describe a crime, including date, number of people killed or injured, incident link, location-related data (latitude, longitude, state, address detail) etc.

We derive new attributes of ***major crime***, which is a categorical indicator that shows whether a crime is a major crime depending on its number of death and injuries.

2. Data Processing

a. For GeoMap, we visualize 500 major crime events out of 260k records to reduce the file size and allow for user interactions. We determine *major crime* with a threshold of 99.8% percentile of the dataset. Data attributes include *incident_id, state, number of death, number of injuries, source_url, latitude and longitude*.

b. For cross filter, we visualize 5k major crime events out of 260k records. We determine *major crime* with a threshold of 98% percentile of the dataset. Data attributes include *state, number of death, number of injuries, total number of death and injuries, and state*.

Goals and Tasks

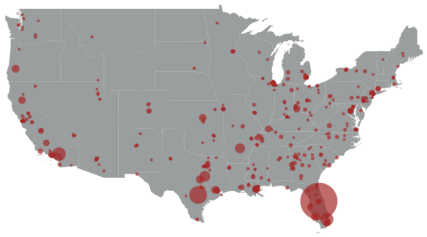
We implement a geographical map regarding gun violence crimes, which includes a hover function, and on-click action. Geographical features are implemented with D3 projection, GeoMap and DOM manipulation methods, the hoover function is implemented with mouse events and tooltip and the on-click function is implemented with mouse events and window open function in React.

We also implement cross filter charts regarding timespan and the distribution of gun violence crime across different states and the number of death and injuries. This feature is implemented with D3 brush and crossfilter library.

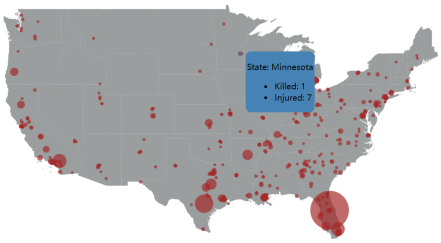
Visualization

There are two main views, first, a U.S map that is filled with points of different colors, and second, two cross-filter bar charts. These two views are linked with each other for analytical purposes as stated above and implemented with hyperlinks.

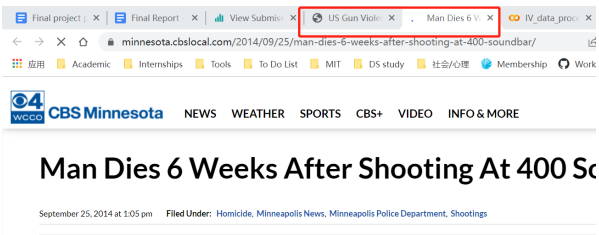
The geographical map contains crime points, each point representing a crime, plotted at the occurrence location.



When users hover over each crime on the map, they will see the color change of this crime. They will find more detailed information regarding this crime on the number of injuries, the number of death and the address of the location of the crime.

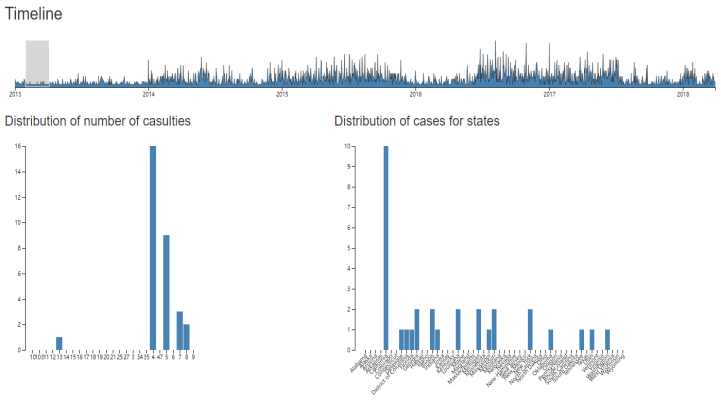


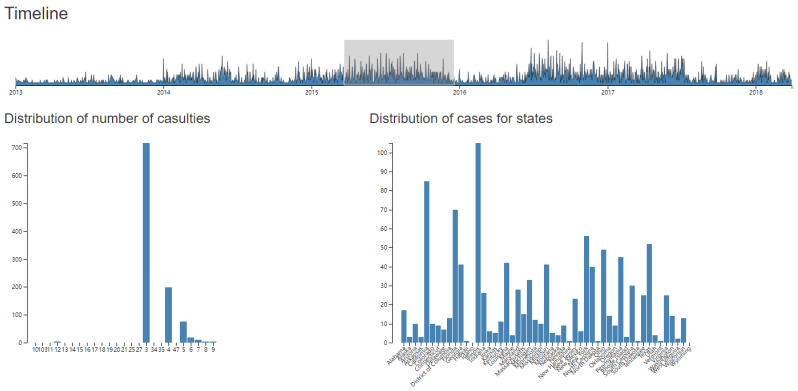
When users click on each crime, they will be directed to the reference link (when links are still traceable) to the reporting source of this crime.



The cross-filter bar charts provide a customized analysis regarding gun violence timespan and according gun violence distribution across states and number of death and injuries. By selecting

a range of time they are concerned with, they can see the difference in the below bar charts and analyze the distribution of gun violence crime across different states.





Reflection

In general, we strictly follow our work plan and timeline we proposed in our proposal for the implementation. While we do change some important points as follows:

In the data processing section, we adjusted the threshold of defining *major crimes* to crop the data for implementation purposes. Our initial plan of 10k data is too large for D3 to function and thus we reduce the data size to 0.5k and 5k data respectively for geomap and cross filter.

In the geomap implementation, we delete the categorized feature of death and injury. Originally, the geographically visualized crimes are divided into two categories, death and injury. However, during implementation, we discover that those crimes represented by bubbles are too close on map to clearly distingusih them without overlapping and causing confusion and therefore delete this visualization. Rather, we add those information as supplementary inside our hover function information box.

In the crossfilter map implementation, we delete the data aggregation part. Originally, we plan to group corresponding data by quarter to retrieve new data column ***date by quarter*** to get a feasible data size for visualization. However, during implementation, this feature adds the number of deaths in different states together and restrains us from getting seperate information of each state, and thus unable to draw the bar chart of the distribution of death regarding states. We turn to focus only on major crime events and use each day as a time unit.