Project Final Report
CSE564 Visualization and Visual Analytics
Unites States Crime Data Visualization

Nitesh Idnani **Rohit Patil**
111986873 **112043953**

22 May 2019

(YouTube video link : https://youtu.be/3bfEIaiPwC4)

1 Introduction

We have developed an interactive data driven visualization of the crime data in the United States over a web platform. The visualizations are designed to help produce certain statistics about the criminal data related to crime distribution throughout the US for various crime types across the past 2 decades, along with percentages of crime in various states and the races affected by those crimes. Information on such trends can help law enforcement agencies and the government devise plans and effective strategies to assist detectives in identifying suspects and hence develop solutions to prevent crimes in future.

2 Dataset

For the visualization of crime statistics and trends, we use the data provided by the UCR[1] which is maintained by the FBI. The site provided us with crime statistics(counts, distribution) for each year starting from 1994. We extend the statistics to produce an extensive consolidated data which provides statistics for each year for various crime types (Burglary, Murder, Violent Crimes, etc) mentioned in more detail in the below section.

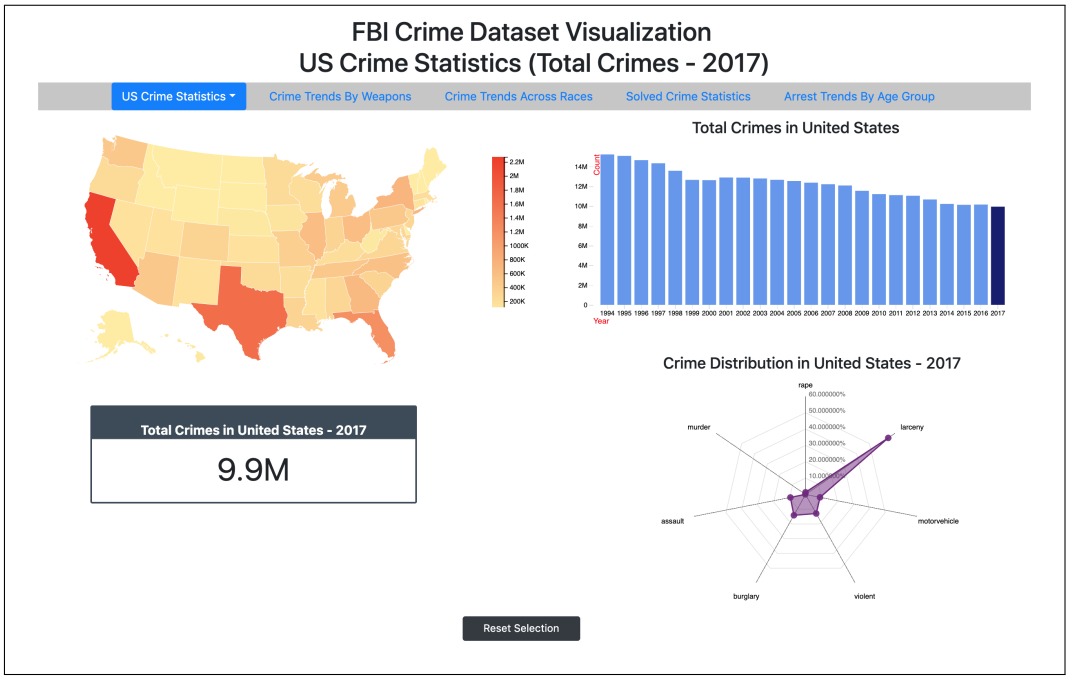
We have used the data to generate certain attribute specific crime stats which provides the necessary data required for visualization of crime trends by weapons, as well as crime trends across various races in the United States. We also extract tables from the site related to the arrest trends for the year 2016-2017 for various age group and solve crime statistics for the past years for different agencies.

3 Visualization and Analysis

The data obtained from the FBI dataset [1] has been categorized and integrated into different data files according to the crime type. As mentioned in the proposal, this data has been processed to make it consistent and free of any missing values (data pre-processing steps have been applied) before rendering data driven visualizations. Moreover, on preliminary analysis of the various attributes from the dataset like murder, larceny, violent crime, etc. we observe that dimensionality reduction technique like PCA and MDS do not provide any relevant insight as there is no particular important crime type that affects a state's condition, rather, an analysis of all the crime distribution in a state and how they affect the population is more relevant and help develop more useful inferences.

We have visualized the data obtained to render the following graphs in this phase in addition to the visualizations created i.e. Crime Trends By Weapons and Races in the previous phase:

1. **US Crime Statistics (Dashboard)**



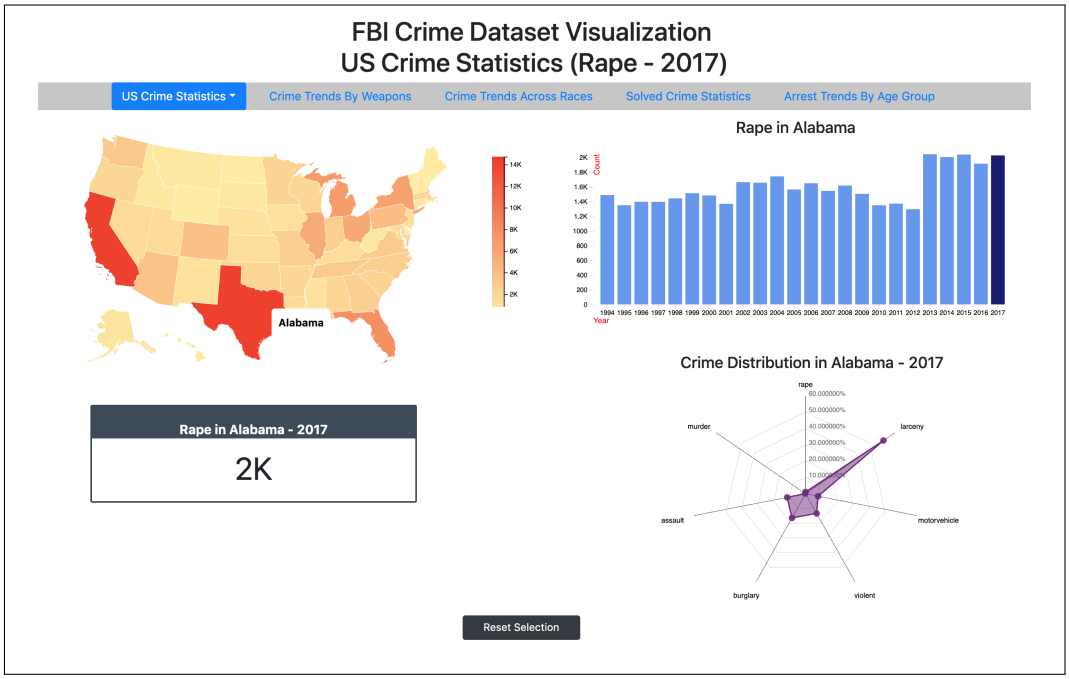
The crime data visualization dashboard consists of three main charts - choropleth map [2] representing the crime numbers in each US state, a bar chart representing the crime numbers over the years (1994 - 2017) and a radar chart [2] representing the distribution of a crime types. As per the dashboard definition learned in class, the most important visualization (chloropleth crime map) is placed in the left hand corner to attract the user's attention. The legend representing the mapping of the crime numbers to color scale is placed right next to it for ease of understanding. The three important variables that can be selected to visualize the crime data are -

(a) US State (50 states)

(b) Year (1994 - 2017)

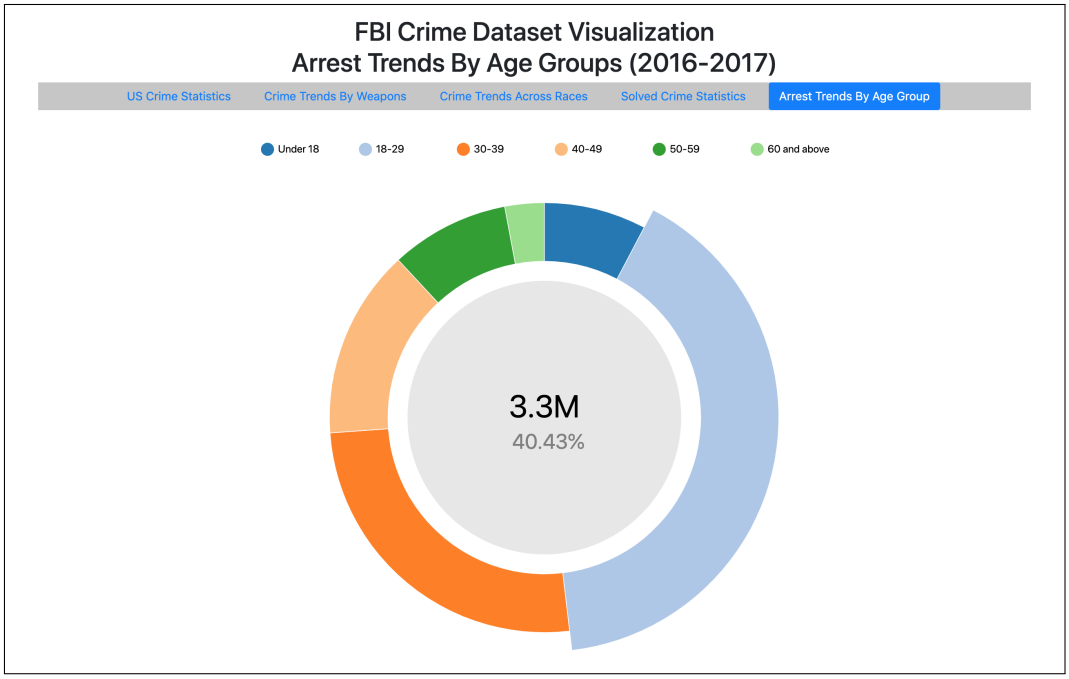
(c) Crime Type

A US state can be selecting (on-click) using the chloropleth map. Similarly, the year can be selected (on-click) using the bar chart and the crime type (like murder, rape, larceny, etc.) using the drop-down in the navigation bar. The default dashboard view is of the total crimes in the United States. The crime numbers of the current selection are displayed in a box (on the lower left) and gets updated depending on the user input (crime type, year). All the charts get updated depending on the crime type, year and location (US state) presenting a coordinated view.



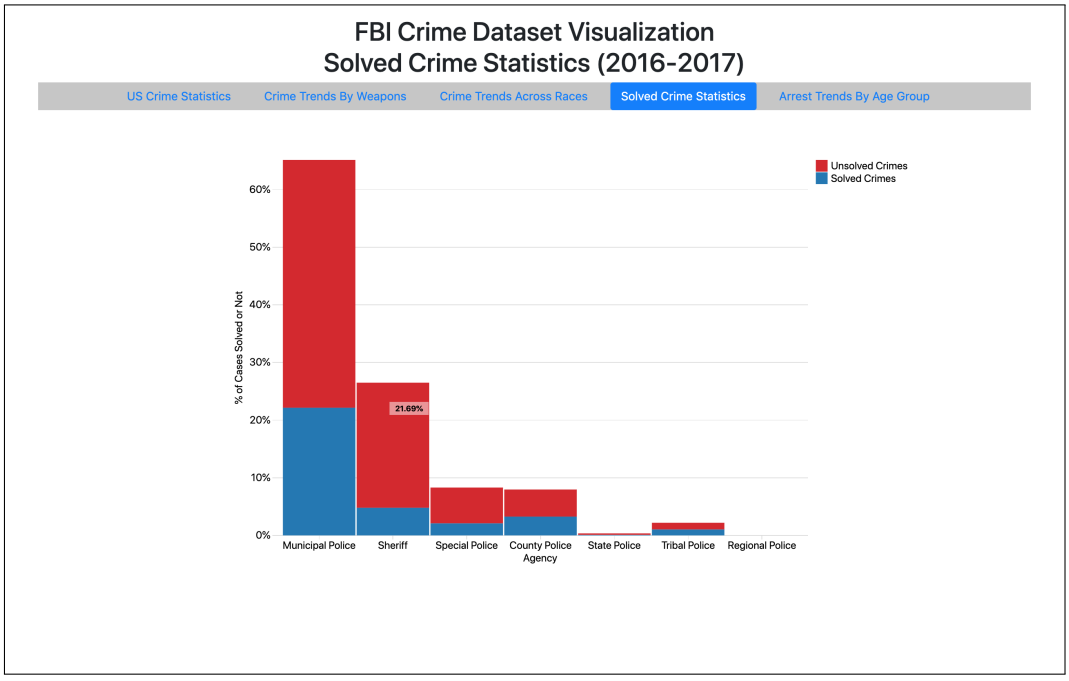
The above dashboard view represents the rape numbers for the year 2017 in the state of 'Alabama'. The state is selected by clicking on the chloropleth map (highlighted using the tool-tip) while the year is selected from the bar graph. The selected year is highlighted in the bar graph. It is easier to identify the crime trends of a particular crime using the bar chart. Even though the rape numbers in Alabama (2K) are less as compared to other states, they have been increasing since the past few years. Such data can help produce valuable insights into the real world (discussed later). Larceny numbers are quite high as compared to rape in the state of Alabama for year 2017. The current selection can be reset using the 'reset selection' button provided on the dashboard. Thus, the dashboard has been designed keeping in mind the principles to convey all the information from multiple datasets to the user. This makes the data easy to interpret and helps develop observations and insights.

2. **US Arrest Trends by Age Group (Donut Chart[2])**



The above visualization shows the overall arrest trends for various age group for the year 2016-2017. We have developed a donut chart with each sector representing the percentage of arrests corresponding to the specific age group. The user can hover over the sectors to see percentage of arrest for the selected age group. Also, the user can select multiple age groups combining one or more age groups and see the combined percentage for the age group in the center.

3. **US Solved Crime Statistics (Stacked Bar Chart [2])**



The visualization provides an account of the extent of solved and unsolved crimes by various government and public agencies. In the above visualization, the red section represents the percentage of crimes that are unsolved and the blue section represents the solved crimes percentage. We can hover over the specific section for any of the agency to see the respective percentage of solved/unsolved crime for that agency.

4 Observations and Insights

- As seen from the dashboard view of the state of Alabama, the rape numbers are on the rise since a past few years. Recently, the state of Alabama passed an abortion ban law which is being contested by many people. [3] The abortion ban holds true even in case of rape or incest. The increasing rape numbers (as seen from dashboard 2) can help to form an argument in favor of those who are against the ban. Data is supportive and it represents facts.

- Similarly, an article published in March 2019 [4] is in accordance with the facts that can be seen from the dashboard visualization of New York. The overall crime is decreasing steadily in New York, and all other states over the past decade. However, the rape numbers have been going up with the highest recorded in the last year. Supporting news with data visualizations can convey the message to the masses and might engage the Agencies in taking appropriate actions.

- California is a state with the highest crime rate according to the chloropleth map color scale mapping. Also, according to the 'crime by weapon' bubble chart, we observe that gun are the most used weapon for deaths in US as cause as many as 5 times more deaths compared to other weapons. The article published by CNBC supports our observations. [5] and our conclusions. It states that stricter gun laws will take into effect from 2019 in the state of California.

- As seen from the donut chart of the arrest trends by age group, we observe that millennials are most likely to commit crimes and get arrested. A news article published by New York Post - 'Young Americans are more likely to be arrested than previous generations' [6] is in accordance with our findings from the data. The result can be used to initiate programs targeted for the specific age group population and hence reduce criminal activities.

- Based on the 'crime trends by race' line chart, we can conclude that Whites people are the most affected by crime activities and Asians are the least affected population group. Also, over the past few years, victimization of Native Americans has been increasing which needs to be checked in future.

- Overall trend in the 'solved crime' stacked bar chart suggests that more than half of the crimes recorded are unsolved for all the agencies in the United States with local governments like Municipal Police receiving the most cases and better at tracking and solving crimes.

5 Implementation

1. **Python, Pandas** - The basic data analysis and pre-processing is done using python and supporting libraries like pandas.
2. **D3.js** - D3.js (Data-Driven Documents) is a JavaScript library for producing dynamic, interactive data visualizations in web browsers. The statistical state-wise data can be visualized via chloropleth maps, line charts for time series comparison of crimes, etc.
3. **HTML, CSS** - This will serve as the client side part for UI styling and improving the user experience.
4. **Python HTTP server** - This will act as the back-end to handle the incoming requests from the client.

6 Conclusion

'Data is beautiful' and can convey a lot if visualized in the right way. We believe that crime data visualization might help us formulate a story for crime reduction or increase in certain geographical locations in the US and make comparisons. Visualizations speak for themselves and can also form the grounds (prove or disprove) of a formulated theory. Our goal is to make the most of the published crime data by visualizing different categories, distributions and percentages and then draw conclusions (story) from them. Similarly, we aim to maximize our learning and apply the concepts learned in the visualization course to create impactful visualizations of the crime data.

References

[1] UCR FBI Crime Dataset
https://ucr.fbi.gov/crime-in-the-u.s

[2] D3 Examples
https://github.com/d3/d3/wiki/Gallery

[3] Alabama Abortion Ban News
https://www.vox.com/policy-and-politics/2019/5/18/18630562/alabama-abortion-ban-voters-exemption-for-rape-incest

[4] Decreasing Overall Crime NY news
https://abc7ny.com/nypd-crime-down-overall-but-murders-and-rapes-up-in-2019/5167247/

[5] California stricter gun laws news
https://www.cnbc.com/2018/12/26/tougher-gun-laws-to-take-effect-in-2019-in-ca-s
html

[6] New York Post - Arrests by Age News
https://nypost.com/2019/02/26/young-americans-are-more-likely-to-be-arrested-tha