**DATS 6401\_12 - Group 1: Analyzing Crime statistics across US**

**Introduction and background.**

Violent crimes have been increasing drastically over the last several years. Crime and public safety are one of the major issues that we must deal with. With the rise in the number of gun violence in the current few years, do we ever wonder where it all started? Was it the same situation 10 years ago? From several articles, we see that the crime rates have specifically increased from the year 2012. Connecticut saw major gun violence towards the end of December in 2012 bringing our attention to mass shootings and mental health. As of 2023, it is considered the deadliest mass shooting at an elementary school in US history. There was also the belief that the Mayan Calendar marked 21st December 2012 as the end of the world. There are several articles that relate this belief with that of the murder-suicide in the Dorchester County in the month of November of 2012. The 2012 crime dataset will be examined in this research to identify trends, demographics, and offense categories. We'll also investigate how much crime there is on a daily basis and look into any connections to issues such as gun violence, racial/ethnic prejudice, and the Mayan Calendar. In order to better understand the social and economic variables that contribute to high crime in some places, we will map crime rates using GIS and identify hotspots. In order to spot trends in criminal behavior, we'll also look at data on crime rates by race and ethnicity. In the end, our research will help shape laws and procedures meant to lessen crime and increase public safety.

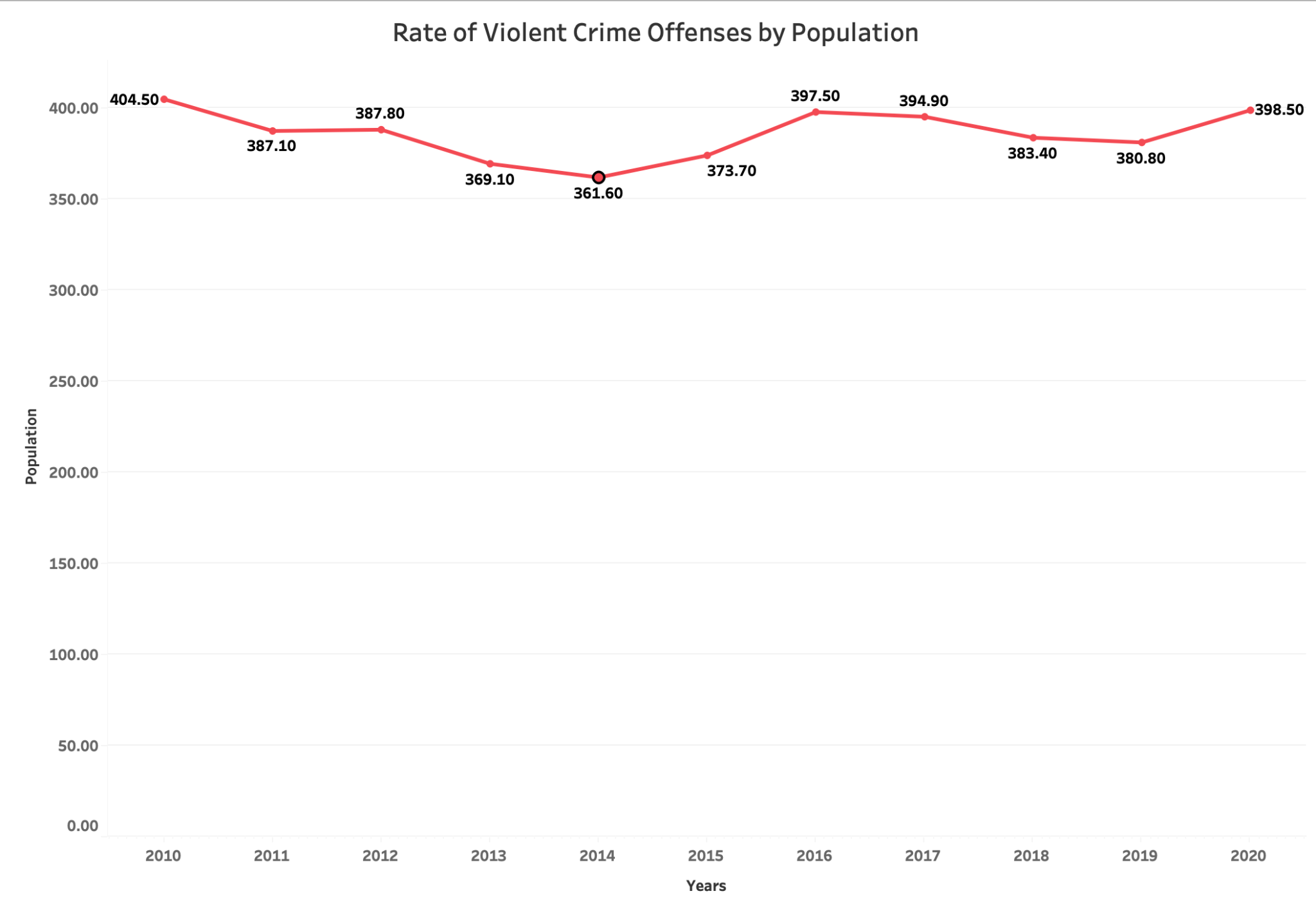
**Datasets:**

The first dataset was collected by the Federal Bureau of Investigation's Uniform Crime Reporting (UCR) Program through the National Incident-Based Reporting System (NIBRS). The UCR Program has been collecting crime data since the 1930s, and NIBRS was implemented in 1989 as a more detailed and comprehensive way to collect crime data. The project is funded by the U.S. Department of Justice, and law enforcement agencies across the United States contribute data to the system. The data was created to provide a comprehensive and detailed picture of crime across the United States. The purpose of the UCR Program is to provide law enforcement agencies, policymakers, and the public with accurate and timely crime statistics to inform decision-making and resource allocation. The dataset in question dates back to 2012 and includes variables related to offenses committed, offenders, victims, agencies reporting the data, and universities reporting data. It contains over 4,000 rows and 63 columns (variables) and includes data for all 50 states in the United States, as well as data for cities within those states. The dataset can help answer questions about crime patterns and trends in the United States, particularly with respect to demographic factors such as race and gender. However, there may be limitations to the data, such as underreporting of certain types of crimes or biases in the way that law enforcement agencies collect and report data. Additionally, the data may not provide a complete picture of the extent and nature of crime in the United States.

The second dataset focuses on the sex and race of the victims by age range, collected by the FBI UCR Program's National Incident-Based Reporting System (NIBRS). The project is funded by the Department of Justice and is a collaborative effort between law enforcement agencies across the United States. The dataset was created to provide a more comprehensive and detailed understanding of crime incidents in the United States. It includes information on victims, offenders, and the context of the crime. The variables in the dataset include the Type of Victim, Age of Victim, Sex of Victim, and Race of Victim. The NIBRS dataset has been collected since 1982 and includes data on over 6 million incidents and 10 million victims. The dataset provides important insights into the distribution of crime victimization by demographic factors, such as age, sex, and race, which can inform policies and interventions aimed at reducing crime and improving public safety. However, there are limitations to the data, including potential biases in reporting and underreporting of certain types of crime, which must be taken into account when interpreting the data.

**Data Story**

1. **Line chart**



The line chart above depicts the rate of violent crime offenses by population from 2012 to 2020. The chart tells us that there was a dip in violent crime offenses by population in 2014 compared to other years, which suggests that efforts to combat violent crime may have been successful during that year. However, the chart also shows a steady increase in the rate of violent crime offenses from 2018 to 2020, indicating that more needs to be done to address the issue of violent crime. The formula for calculating the crime rate is also presented, which helps the audience understand how the data was calculated. Overall, this chart provides important insights into the trends and patterns of violent crime offenses by population over the years, which can inform policies and interventions aimed at reducing violent crime and improving public safety.

Crime rate = (Number of reported crimes / Population) x 100,000

1. **Table**

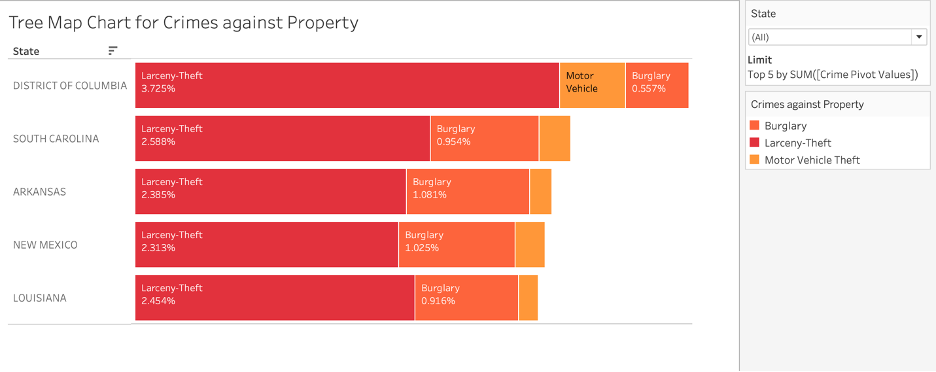


The data provided is about the number of deaths in the United States by race and age group. The data shows that the majority of deaths are among white individuals, with a total of 2,938,479 deaths. The next highest group is Black individuals, with a total of 847,453 deaths. The third highest group is Asian/Pacific Islander individuals, with a total of 46,933 deaths. American Indian/Alaskan Native individuals have the lowest number of deaths, with a total of 20,975 deaths.

When looking at age groups, the data shows that the highest number of deaths occur in the 66 and over age group, with a total of 474,045 deaths. The next highest group is the 56-60 age group, with a total of 226,740 deaths. The lowest number of deaths occur in the 10 and under age group, with a total of 72,828 deaths.

The data is broken down further by race within each age group. Within each age group, the highest number of deaths occur among White individuals. The second highest group varies by age group, with Black individuals being the second highest in most age groups, except for the 10 and under age group where the second highest group is individuals of unknown race.

1. **Tree Map**



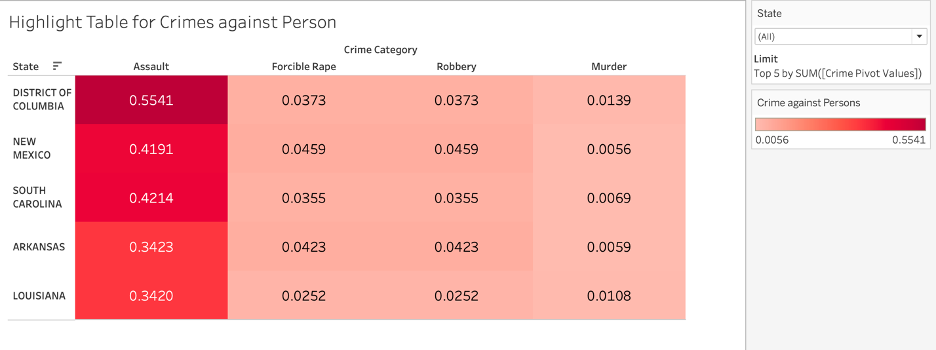
Our next objective is to identify the top 5 states in the USA with the highest crime against property rates. This category of criminal offenses involves the illegal taking, damage, or destruction of someone else's property without causing harm or injury to anyone else. We specifically investigate three offenses: Burglary, Larceny-theft, and Motor Vehicle Theft.

With the help of tree maps, we observe that larceny theft has a slightly higher rate than the other two offenses. This is because larceny theft is a more general term that covers thefts of personal property, including shoplifting, pickpocketing, and thefts from vehicles. In contrast, burglary involves entering a building or structure with the intent to commit a crime, and motor vehicle theft involves stealing a vehicle, which may be more difficult to carry out.

Once again, DC has the highest crime against property rate, likely due to its status as the capital, which attracts many tourists and visitors. Additionally, high income inequality and poverty rates in some areas of DC may contribute to higher rates of property crimes.

It is important to note that DC, New Mexico, South Carolina, Louisiana, and Arkansas appear in the top five for both crime against person and crime against property charts. These states have higher poverty rates, lower education rates, and higher unemployment rates than other states, contributing to their higher crime rates. However, it is crucial to recognize that the reasons for high crime rates are complex and can depend on various factors such as social and economic conditions, demographic makeup, and historical trends.

1. **Highlight Table**

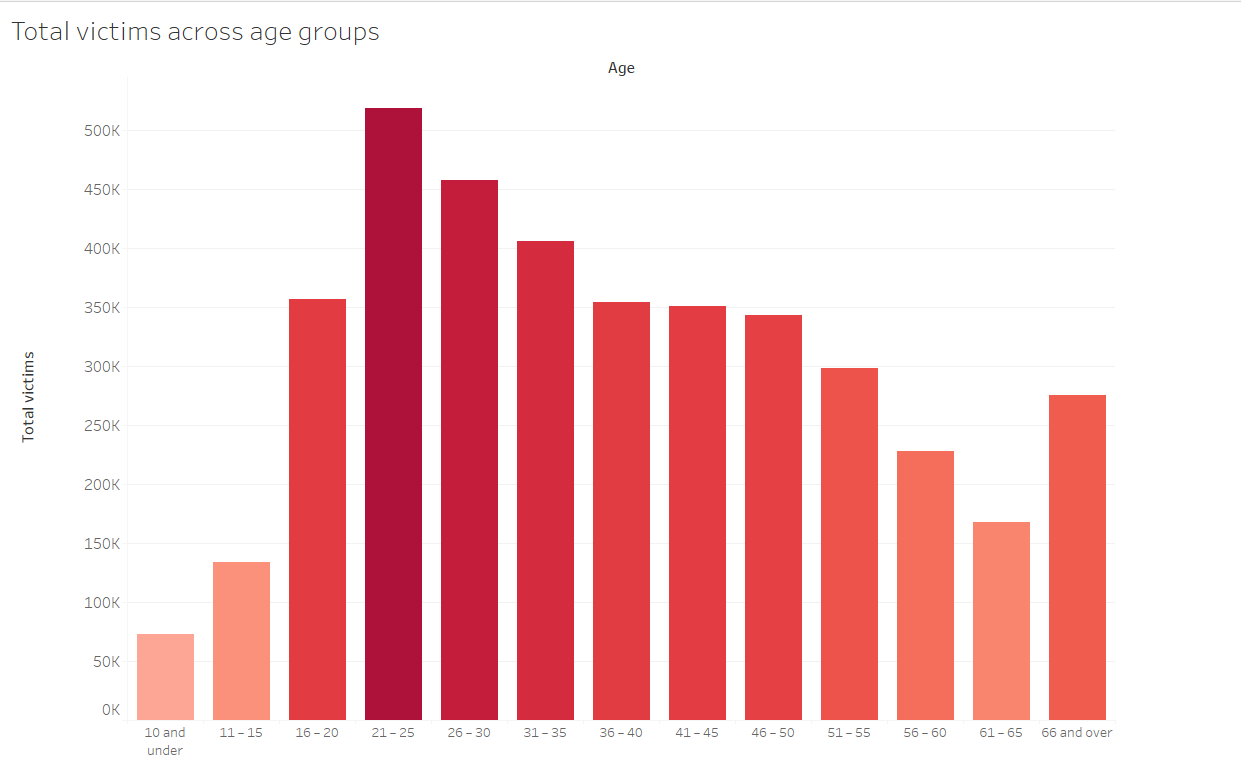


We have created a highlight table using table 5 of the FBI dataset to identify the top 5 states in the US with the highest crimes against persons - Assault, Forcible rape, Robbery, and Murder. Crimes against persons are criminal offenses that involve physical harm or injury to individuals. Each offense is a calculated field, representing the specific offense divided by the population of that state to get the crime rate per capita.

Notably, for each state, the number of Assault crimes is always higher than those of murder, rape, and robbery. This could be due to several factors such as the availability of weapons, social or economic factors, and cultural attitudes towards violence.

Surprisingly, the District of Columbia, the US capital, has the highest crime rate. This could be due to the high population density in DC, which provides more opportunities for criminal activities to occur. Additionally, the presence of government buildings and agencies in DC may also make it a target for terrorist activity, which can also contribute to a higher crime rate.

1. **Bar chart**

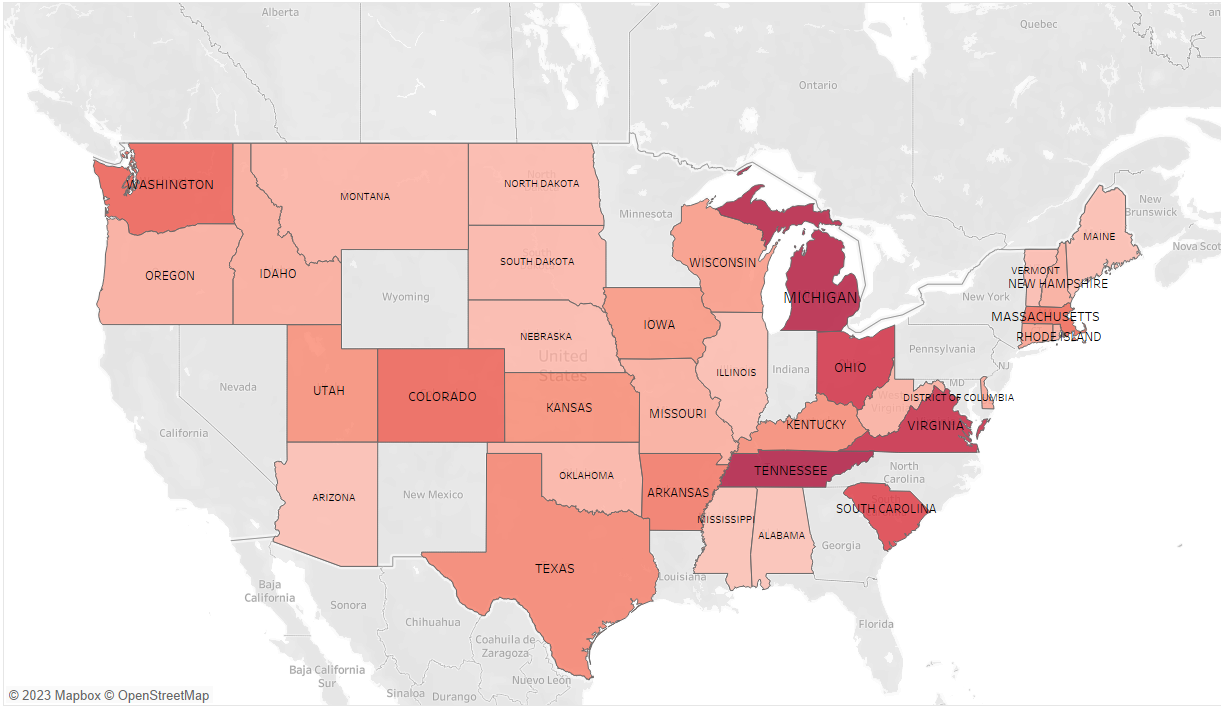


The bar chart represents the number of total victims under different age groups for the year 2012. The dataset chosen for the above graph is from the National Incident based reporting system website which consists of columns like age, sex, race of the victims. We can see from the above bar chart that the number of victims are the highest for the age groups 21-25 and lowest for 10 and under.

The dataset does have few null values and unknown values for the age column which is a limitation. These unknown values are also present for other columns like sex and race. Another limitation of the dataset is that, for the age groups 66 and over it shows a higher value for total number of victims but we really don’t know how it varies for ages greater than 70 or less than 90. We can only find the total number of victims for 66 and above.

We can also observe that the total number of victims is 72,828 for age group 10 and under which only keeps increasing up to 5,18,632 for the age group 21-25. After which we see that there is a continuous decline in the total number of victims to 1,67,797 until the age groups 61-65. However, at the end there is another increment to 2,75,070 for the age groups 66 and over.

1. **Geographical Map Chart**



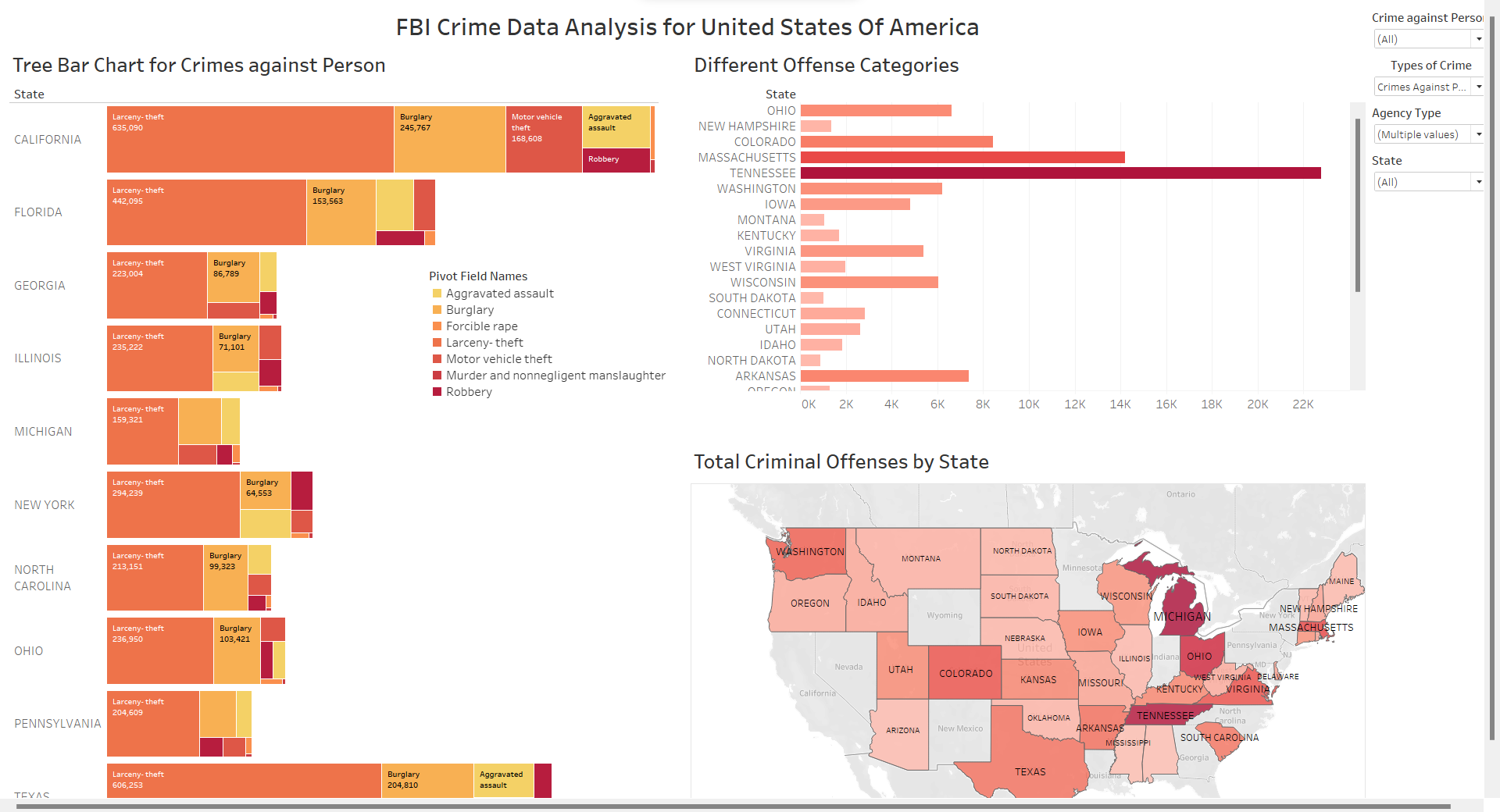
The geographical map is designed to provide an overview of the number of recorded offenses in each state in the US for the year 2012(Variables used: Total offenses, state, Crime against person, Crime against property, and crime against society). The map shows the total number of offenses in each state. By looking at the map, viewers can see which states have the highest number of offenses and get an idea of how different regions were impacted by crime that year. However, the map also includes additional information for each state, such as details on the types of crimes committed against persons, property, and society, as well as the state's population. This allows viewers to dig deeper into the data and identify specific offenses for each state, which can be useful for identifying trends or patterns in crime rates.

Overall, this map is useful for understanding the distribution of offenses across the US in 2012, and provides a starting point for further analysis and investigation into the factors that may have contributed to crime rates in different regions. From this chart, we can observe that Michigan, Tennessee, and Virginia have the highest number of offenses in the year 2012.

One important thing to note about this dataset is that data for several states, including California, Nevada, New Mexico, Florida, and North Carolina, as well as a few others, are not available. This means that the visualization may not provide a complete picture of crime rates across the US, and caution should be exercised when drawing conclusions or making comparisons between states with and without data.

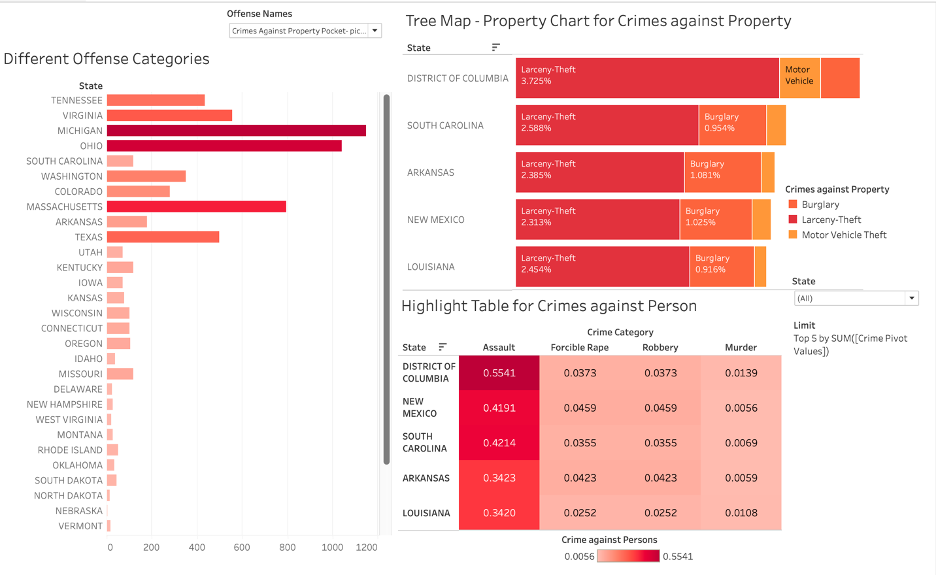
**Dashboard Screenshots**

1. **DASHBOARD 1**



The Dashboard above shows the crime analysis in the United States based on the FBI data. Using the "crime against the person" filter, the left chart allows for a deep dive analysis of crimes against individuals. Users can further refine their search by crime type (which allows only one selection at a time as it is easier for the user to have a deeper analysis on different types of crime), agency type (including city, metropolitan, nonmetropolitan, state police, universities and colleges, and other agencies), and state. These filters have an impact on the charts on the right side of the dashboard, which shows information on various offense categories as well as the total number of criminal offenses by state.

1. **DASHBOARD 2**



The displayed dashboard presents offense categories and their occurrences in different states. The left-hand bar chart enables a comparison of the specific offense category among all the states. With over 50 offenses, the chart provides an extensive view of the data. On the right-hand side, there are two charts that allow for a deeper exploration of crime against persons and crime against property. The displayed chart exhibits only the top 5 states with the highest crime rate in each domain. However, the interactive feature of the dashboard enables the selection of one state at a time, providing a detailed examination of its crime rate numbers.

**Summary and Conclusions**

The visual representations offered offer significant insights into the trends of crime and mortality in the United States. The tree bar chart and highlight table, specifically for Crime against Person and Property, reveal that California has the highest figures across all categories, with Assault and Larceny theft exhibiting the highest rates, respectively.

The line chart on violent crime offenses reveals a dip in offenses in 2014 but a steady increase from 2018 to 2020.

Moreover, the table on deaths by race and age group indicates that white individuals have the highest number of deaths, and the 66 and over age group has the highest number of deaths. The tree map on crimes against persons identifies the top 10 states with the highest offenses, with California being the most concerning. The bar chart on total victims by age group shows that the 21-25 age group has the highest number of victims.

However, these visualizations have limitations, such as null values and the inability to determine the total number of victims for ages beyond 66. Despite these limitations, policymakers and researchers can utilize the insights provided by these visualizations to develop effective strategies to address crime and mortality trends in the US. For example, addressing the steady increase in violent crime offenses, particularly in California, and targeting interventions to reduce crime rates in the 21-25 age group could be prioritized.

In summary, the visualizations highlight the severity of crime and mortality trends in the US, particularly in California. They provide an opportunity to inform evidence-based policymaking and further research to address these issues.

**References:**

**1. FBI. (2012). Data tables: National Incident-Based Reporting System. Retrieved from** [**https://ucr.fbi.gov/nibrs/2012/data-tables**](https://ucr.fbi.gov/nibrs/2012/data-tables)

**2. Federal Bureau of Investigation. (n.d.). Crime data explorer: Crime trend. Retrieved from** [**https://cde.ucr.cjis.gov/LATEST/webapp/#/pages/explorer/crime/crime-trend**](https://cde.ucr.cjis.gov/LATEST/webapp/#/pages/explorer/crime/crime-trend)

**3. Bureau of Justice Statistics. (n.d.). National Incident-Based Reporting System (NIBRS). Retrieved from** [**https://bjs.ojp.gov/national-incident-based-reporting-system-nibrs#:~:text=NIBRS%20data%20more%20accurately%20reflect,of%20crime%20and%20public%20safety**](https://bjs.ojp.gov/national-incident-based-reporting-system-nibrs#:~:text=NIBRS%20data%20more%20accurately%20reflect,of%20crime%20and%20public%20safety)

**4. Wikipedia contributors. (2023, March 2). Sandy Hook Elementary School shooting. In Wikipedia. Retrieved March 4, 2023, from** [**https://en.wikipedia.org/wiki/Sandy\_Hook\_Elementary\_School\_shooting**](https://en.wikipedia.org/wiki/Sandy_Hook_Elementary_School_shooting)

**5. Domm, A. (2012, December 22). Was Mayan calendar tied to murder-suicide in Dorchester County? Post and Courier. Retrieved from** [**https://www.postandcourier.com/archives/was-mayan-calendar-tied-to-murder-suicide-in-dorchester-county/article\_d47919fd-b473-51e2-981e-df3fcbfa94da.html**](https://www.postandcourier.com/archives/was-mayan-calendar-tied-to-murder-suicide-in-dorchester-county/article_d47919fd-b473-51e2-981e-df3fcbfa94da.html)

**6.** **Florida Department of Law Enforcement. (n.d.). Formulas commonly used in UCR reports fdle.state.fl.us.(n.d.). Retrieved March 6, 2023, from**[**https://www.fdle.state.fl.us/CJAB/Documents/UCR/Formulas-commonly-used-in-UCR-Reports**](https://www.fdle.state.fl.us/CJAB/Documents/UCR/Formulas-commonly-used-in-UCR-Reports)[**.**](https://www.fdle.state.fl.us/CJAB/Documents/UCR/Formulas-commonly-used-in-UCR-Reports.)

**7.** **FBI table 5 dataset tables retreived from Uniform Crime Reporting program**<https://ucr.fbi.gov/crime-in-the-u.s/2012/crime-in-the-u.s.-2012/tables/5tabledatadecpdf>

**Contributions**

1. **Abhishek:** Correlation Matrix Chart Type description, Alternative Chart type suggestions
2. **Akhil:** Stacked Bar Chart Type, Description, References and Citations
3. **Anushka:** Introduction, objectives, goals, Dataset brief description, Tree bar chart for crimes against Property, Highlight table for crime against Persons, Dashboard crime offenses creation
4. **Renganathan:** Bar Chart for analysis of crimes, Geographical Map Chart , Bar graph for total offenses, Dashboard Crime analysis creation.
5. **Sowmya:** Datasets description, Line Chart, Table, Summary and Conclusion, Adding APA formats