Neshat Jalali Heravi November 2021
Business Fundamentals – Final Project

Data Analysis of Gun Violence in America

Background

Gun laws are referred to laws and regulations related to manufacturing, trading, possession, transporting, record keeping, and destruction of firearms, firearms accessories, and ammunition. These laws are enforced by federal Bureau of Alcohol, Tobacco, Firearms and Explosives (ATF) and the state agencies. The objective of this project is to perform Exploratory Data Analysis on gun violence related data to provide insights on this subject matter.

Design & Solution Path

This project is designed to perform Exploratory Data Analysis on gun violence related data and provide insights on this subject matter. Hopefully, ATF and state agencies involved in gun laws can use this information to make more data-driven decisions in proposing and enforcing gun regulations. More effective gun laws can impact thousands of lives of the citizens and law enforcements.

The Exploratory Data Analysis includes US international ranking in civilian arm ownerships, and gun violence in each state in the US. It will also include data and analysis of US mass shootings and demographics of the offender. The impact hypothesis is that by understanding the gun violence data in the US, government agencies can better allocate their resources for enforcing current gun laws, introduce new laws such as imposing a background check on firearms buyers, thus reducing the number of gun violence incidents and casualties. States such as California and Illinois can use this analysis to jump-start projects to find effective solution in reducing number of gun violence incidents. The desired impact would be to reduce the number of gun violence incidents by 10% in cities such as San Francisco or Chicago by allocating more resources based on this analysis.

The data science solution path is to use Poisson regression modeling to predict the number of casualties in a gun violence incident based on location, offenders' demography, city events, number of law enforcements in the vicinity, type of arms involved etc. Also, we can use the unsupervised clustering to segment state counties and locations that have higher rates of gun violence in more detail for better allocation of resources.

Data

Data related to gun violence and mass shootings in the US are obtained from data.world which is an open-source database for this analysis. These three datasets include different countries rankings in gun ownership for 2016, some location data of the mass shootings between 1982 – 2021, and demographic information of the offender and number of casualties etc. Due to the nature of the problem, and some offenders being minors, some details and information are not publicly published.

Tools and Algorithms

For this project, Google Sheet was used as the main tool for data cleaning and manipulations. In addition, SQL was used for minor aggregation of data since Google Sheets slow down for larger datasets. Moreover, Tableau is used for creating visualization and dashboard. The interactive visualization in Tableau allows a more thorough storytelling for the stakeholders.

Communication

You can find the presentation slides and details on the analysis on my GitHub. In addition, the Tableau dashboard includes some interactive visualization of the results.