Comparative Analysis of Firearm Ownership and Crime Rates Across U.S. States

I.   INTRODUCTION

Firearm ownership is always a controversial topic in the United States, where the right to bear arms is protected by the Constitution. However, individual states vary greatly in the regulations they impose on firearm ownership. This project analyses the correlation between firearm ownership, regulation strictness, and crime rates across various U.S. states. States like **"California", "Massachusetts" and "New York"** have strict regulations, while "Texas", "Alaska" and "Wyoming" are more lenient.

By comparing data from both types of states, we aim to identify patterns and assess how regulation differences impact crime rates, offering insights into the relationship between gun ownership and crime. In other words, this project seeks to answer:

"How does the level of firearm ownership and stringency of restrictions correlate with crime rates across different states in the United States, and what differences can be observed between states with stricter vs. more lenient gun laws?"

II.   DATA SOURCES

This report utilizes three primary datasets: FBI NICS Firearm Background Checks [1], Firearm Mortality by State [2], and FBI Crime Data [3]. These comprehensive datasets provide valuable information on firearm ownership granted permits, crime statistics, and firearm mortality rates at the state level. The analysis focuses on six U.S. states, representing two distinct regulatory approaches to firearm ownership. California, Massachusetts, and New York are examined as states with strict regulations, while Texas, Alaska, and Wyoming are analysed as states with more lenient policies.

Given that the raw data are obtained from official organizations, an Extract, Transform, Load (ETL) pipeline is implemented to ensure data quality and facilitate thorough analysis. The ETL process involves several key steps to refine the datasets. These include filtering data to the six specific states of interest, removing irrelevant columns, performing data type

conversions for consistency and grouping and summarizing the data to provide meaningful insights at the state level. Finally, the data are grouped and summarized to provide meaningful insights at the state level. Finally, the cleaned datasets are saved in two formats: as CSV files for easy access and loaded into an SQL database to support more complex queries and future analysis.

A. Data Structure

FBI NICS Firearm Background Checks represents a key component in determining the strictness of firearm-related regulations across states. It provides data on the number of firearm background checks initiated through the National Instant Criminal Background Check System (NICS), covering the period from 1998 to 2023.

The dataset's structure includes both temporal and categorical features, with "year" capturing the trend over time for each specific "state". Besides, it contains several integral features such as "permit", "handgun", "long_gun", "multiple", and "totals", which represent different types of firearm permissions granted after passing the background check. The "totals" feature accounts for the overall number of successful checks, with return cases and failed permits already deducted.

By examining the variations in different types of firearm permits and the overall totals, it becomes possible to draw insights into the strictness of each state's firearm regulations. (*Fig. 1*)

	year	state	permit	handgun	long_gun	multiple	totals
0	1998	alaska	4.0	1365.0	2555.0	61	3985
1	1998	california	5366.0	28641.0	33438.0	0	67445
2	1998	massachusetts	0.0	321.0	1116.0	6	1443
3	1998	new york	1068.0	1094.0	12513.0	12	14736
4	1998	texas	5324.0	23862.0	52248.0	963	82399
5	1998	wyoming	121.0	970.0	2352.0	43	3486

	YEAR	STATE	RATE	DEATHS
0	2022	AK	22.4	164
1	2022	CA	8.6	3484
2	2022	MA	3.7	263
3	2022	NY	5.3	1044
4	2022	TX	15.3	4630
5	2022	WY	20.4	124

Fig. 1. FBI NICS Firearm Background Checks Dataset (left)
Fig. 2. Firearm Mortality by State Dataset (right)

Firearm Mortality by State [2] and *FBI Crime Data* [3] offers crucial insights into the outcomes associated with firearm ownership, serving as a key factor in establishing cause-outcome relationships in subsequent analyses.

Firearm Mortality by State [2] presents comprehensive firearm mortality statistics at the state level across the United States, providing annual data on deaths resulting from firearm-related incidents.

The structure of this dataset contains both temporal and categorical features. The "year" variable captures trends over time for each specific "state", allowing for longitudinal analysis. Besides, the dataset includes continuous features such as "rate" and "death", which respectively represent the proportion of firearm-related deaths relative to the state's population and the actual number of death cases. (*Fig. 2*)

	year	state_abbr	state_name	population	violent_crime	homicide	aggravated_assault	property_crime	totals
0	1979	AK	alaska	406000	1994	54	1203	23193	26444
1	1979	CA	california	22696000	184087	2952	93129	1511021	1791189
2	1979	MA	massachusetts	5769000	30650	212	17286	310756	358904
3	1979	NY	new york	17649000	161906	2092	60949	933234	1158181
4	1979	TX	texas	13385000	67988	2235	34043	725109	829375
5	1979	WY	wyoming	450000	1579	41	1224	20129	22973

Fig. 3. FBI Crime Dataset

Similar to the above datasets for "year", "state_abbr" and "state name", the structure of *FBI Crime Data* [3] also includes integral features, covering population of the state, firarm-related crimes("violent_crime", "homicide", "aggravated_assault", "property_crime") and the total crimes ("totals"). (*Fig. 3*)

The structure of the FBI Crime Data dataset shares similarities with the previously mentioned datasets, incorporating temporal and geographical features ("year", "state_abbr","state name"). However, this dataset extends its scope to include a range of integral features that contains the population of each state and firearm-related crime categories("violent_crime", "homicide", "aggravated_assault", "property_crime").

Therefore, the three datasets collectively provide a comprehensive foundation for analyzing the cause-outcome relationships. They cover key aspects ranging from state-specific characteristics, Firearm Background Checks, and population data to firearm-related crimes and mortality rates. By integrating these diverse datasets, the analysis can explore the correlation and understanding between gun laws, ownership patterns, and public safety outcomes.

B. Licenses

All datasets used in this analysis are obtained from official organizations, ensuring their reliability and authenticity. The *Firearm Mortality by State* is licensed under Section 308(d) of the Public Health Service Act and CIPSEA. The *FBI NICS Firearm Background Checks* and *FBI Crime Data* are licensed under MIT License and accessible through the FBI's FOIA Library respectively. Data is used for transparency, academic research and statistical purposes only. Appropriate credit and link to the license will be provided.
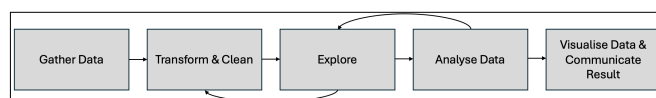
III.   DATA ANALYSIS

Fig. 4. Data Analysis Process

Another thing many data analysts do (alongside cleaning data) is to carry out an **exploratory analysis.** This **helps identify initial trends and characteristics, and can even refine your hypothesis**.

Methodology: XXXXX

Results: XXXXX

Interpretation: XXXXX

IV.   CONCLUSION
A. Summary

B. Limitation and Future Work

REFERENCES

[1]

[2]

[3]