**Crime Data Analysis Proposal**

**By: Luke Truesdale and Michael Hebbeln**

1. **Introduction**

Crime trends offer critical insights into the safety and stability of communities, but what factors drive these patterns? Are economic conditions, such as income levels and unemployment, or political leadership key influences on crime rates? In this project, we will use a 2018 crime dataset sourced from Kaggle to analyze crime patterns across U.S. communities. To enrich this dataset, we’ll supplement it with unemployment rates from BLS.gov, and mayoral data from USMayors.org and Ballotpedia.org, which includes the political affiliations of local leaders. This combined approach will allow us to explore how economic conditions and local governance shape public safety.

1. **Data**

The first data source for this project is a crime dataset from Kaggle, which contains detailed 2018 data on U.S. communities. It includes both demographic features such as population size, race composition, and household size, and crime statistics such as burglary incidents, larcenies, arsons, and violent crime rates per capita. This dataset forms the foundation of the analysis, offering insights into the distribution and frequency of crime across various communities.

We will incorporate unemployment data from the Bureau of Labor Statistics (BLS). Economic instability, represented by high unemployment rates, is often linked with higher property crime rates. Using BLS APIs, we will collect city-level or regional unemployment data. If direct matches for specific communities are not available, we will link unemployment data to the closest major city, county, or state if need be to ensure accurate analysis.

Additionally, we scraped mayoral data from USMayors.org, which provides the names of mayors from the top 1400 major U.S. cities. We also scraped political affiliations of these mayors for the top 100 largest cities from BallotPedia.org. Political leadership can influence public safety policies and law enforcement priorities.

<https://www.kaggle.com/datasets/michaelbryantds/crimedata>

<https://www.bls.gov/>

<https://www.usmayors.org/>

<https://www.ballotpedia.org/>

**Data Dictionary:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Column** | **Type** | **Source** | **Description** |
| communityName | Text | crimedata.csv | Name of the community |
| state | Text | crimedata.csv | state where the community is located |
| population | Numeric | crimedata.csv | total population of the community |
| householdsize | Float | crimedata.csv | average size of households in the community |
| racepctblack | Float | crimedata.csv | Percentage of Black residents |
| racePctWhite | Float | crimedata.csv | Percentage of White residents |
| racePctAsian | Float | crimedata.csv | Percentage of Asian residents |
| racePctHisp | Float | crimedata.csv | Percentage of Hispanic residents |
| bulgaries | Numeric | crimedata.csv | Number of burglary incidents |
| larcenies | Numeric | crimedata.csv | Number of larceny incidents |
| arsons | Numeric | crimedata.csv | Number of arson incidents |
| ViolentCrimesPerPop | Float | crimedata.csv | Violent crimes per capita |
| HouseholdIncome | Float | statistia.com | Avg Household income of the community |
| Mayor of Community | Text | usmayors.org | Mayor of largest city by population in community |
| PoliticalLeaningofMayor | Text | usmayors.org | Political leaning of mayor |
| Unemployment Rate | Float | bls.gov | Unemployment rate of largest city by population in community |

1. **Analysis**

This project aims to explore the relationship between crime rates, socioeconomic conditions, and political leadership across U.S. communities. By enriching the crime dataset with household income, unemployment rates, and mayoral data, we will investigate how economic and governance factors shape public safety. Specifically, our research questions include:

* Is there a strong correlation between economic conditions, such as household income or unemployment, and crime rates? Do communities with higher income levels experience fewer property crimes?
* How does political leadership influence public safety? Are communities governed by mayors with specific political leanings more likely to experience different crime patterns?
* Are there regional differences in crime rates? Which communities within the dataset show the highest levels of crime, and do these align with socioeconomic trends?
* What types of crime are most influenced by economic instability? Specifically, does high unemployment correlate more with property crime than with violent crime?

*3.1 Economic Conditions and Crime Rates*

**Visualization:** Scatterplots of householdincome vs. violentcrimesperpop and pctunemployed vs. violentcrimesperpop.

**A graph with a red line and blue dots

Description automatically generated**

**Description:** The analysis explored the relationship between household income and unemployment rates with violent crime rates. A scatterplot of householdincome vs. violentcrimesperpop shows a moderate negative correlation (-0.45), indicating that higher income levels are associated with lower violent crime rates. Similarly, unemployment rates (pctunemployed) exhibit a weaker positive correlation (0.22) with violent crime rates, suggesting that higher unemployment may slightly increase violent crime. These findings align with the hypothesis that economic stability reduces crime prevalence, though other factors likely contribute.

*3.2 Political Leadership and Crime*

**Visualization:** Bar chart of violentcrimesperpop grouped by politicalleaningofmayor.

A graph of a graph with red and blue dots

Description automatically generated

**Description:** The bar chart compares average violent crime rates for communities with different mayoral political leanings. Communities led by mayors with liberal affiliations exhibited higher average violent crime rates compared to conservative or independent affiliations. This difference could reflect urban-rural divides, as larger cities tend to have higher crime rates and liberal leadership. However, causation cannot be inferred without further analysis. This finding underscores the need to consider geographic and demographic factors in interpreting the impact of political leadership on crime.

*3.3 Regional Trends in Crime Rates*

**Visualization:** Bar chart of violentcrimesperpop by state.

A graph of a number of blue dots

Description automatically generated

**A graph of a crime

Description automatically generated**

**A black text on a white background

Description automatically generated**

**Description:** The state-wise bar chart highlights significant regional variations in violent crime rates. States like DC and South Carolina report the highest violent crime rates, while states such as Maine and Vermont exhibit the lowest. The top five states with the highest crime rates are concentrated in the South and urbanized regions, indicating potential links to socioeconomic factors or regional disparities in law enforcement policies. These findings emphasize the importance of tailoring crime prevention strategies to regional needs.

*3.4 Types of Crime Influenced by Economic Instability*

**Visualization:** Scatterplots with regression lines for pctunemployed vs. individual crime types (larcenies, burglaries, arsons).

**A graph of unemployment rate and regression line

Description automatically generated**

**A graph of unemployment rate and a red line

Description automatically generated**

**A graph of unemployment rate and regression line

Description automatically generated**

**Description:** The scatterplots examine unemployment rates and their relationship with property crime categories. The strongest correlation is observed between unemployment and burglaries (0.35), followed by larcenies (0.28), and a weak correlation for arsons (0.12). Regression lines indicate a positive relationship across all crime types, suggesting that economic instability exacerbates property crimes. These insights support policies aimed at economic stabilization as a mechanism for crime reduction.

1. **Conclusion**