Report on Changes in Crime Rates in Major U.S. Cities

Introduction

The objective of this project was to analyze changes in crime rates in major U.S. cities over the past 40 years. The research question for this project was, "How has crime changed over the past 40 years in U.S. major cities?" This topic was chosen as it was deemed interesting and potentially informative to analyze crime changes in major US cities. Additionally, there was a significant amount of data available that would allow for the creation of a meaningful and clear project.

Methods

To analyze the data, a k-means clustering algorithm was used to group cities into low, mid, and high crime clusters based on their homicide rates per 100,000 people. Data was collected from 1975, 1995, and 2014. To ensure that the data was standardized, any NaN values were dropped, and only data that was reported over a 12-month period was analyzed. This decision was made to standardize the data and to eliminate any potential outliers that may have skewed the results.

Results

The results of the analysis indicate that the majority of cities have shifted from higher crime groups to lower crime groups over the period of 1975, 1995, and 2014. While there were outliers such as New Orleans that saw an increase in crime rates, the overall trend was a decrease in crime rates.

Interactive map visualizations were used to communicate the data, with each color representing a different crime cluster group (low, mid, and high). The redder the color of the city, the more homicides occurred.

Figure 1 - Geo Heat Map Comparing Homicides per 100k in Major Cities Across the United States in 1975

Figure 1 shows that the homicide rate varied between 1.64 and 44.19 per 100k in 1975**.** While there are many reddish bubbles, these are on the lighter side, indicating relatively low rates nationwide.
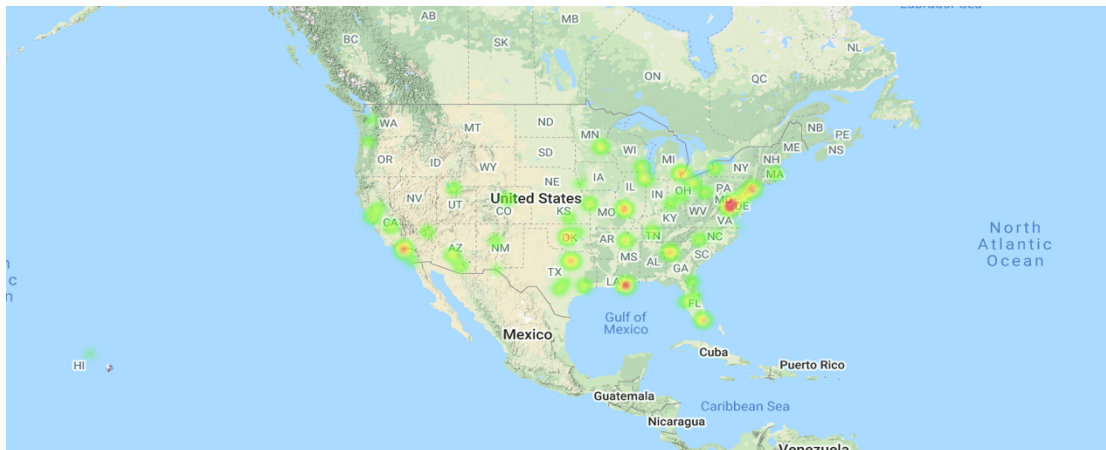
Figure 2 - Geo Heat Map Comparing Homicides per 100k in Major Cities Across the United States in 1995

In 1995 homicide rate varied between 1.66 and 74.51 per 100k. This was the highest rates of all three years and indicates an uptick in crime and homicide from 1975. In this case, while fewer red bubbles than 1975, these are of greater intensity thus indicating higher homicide rates.
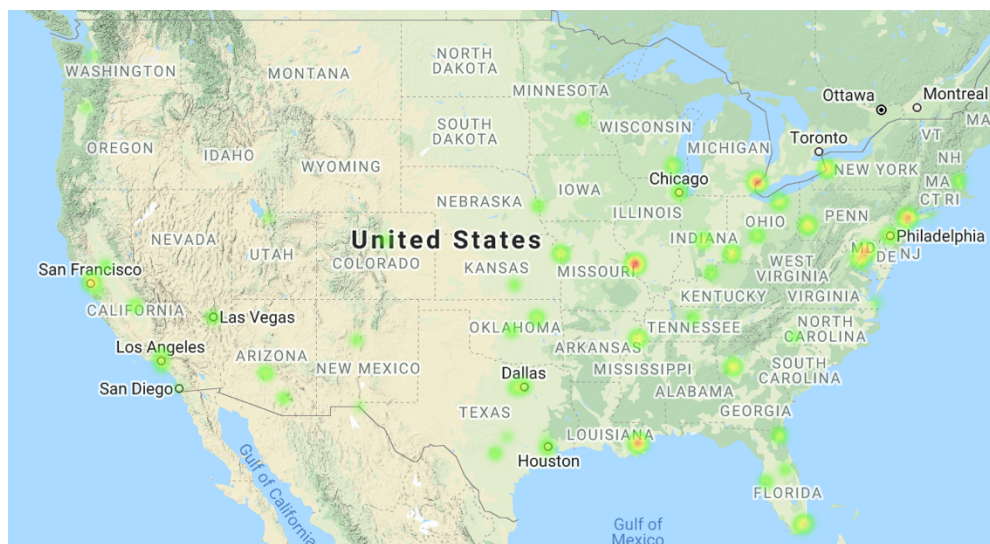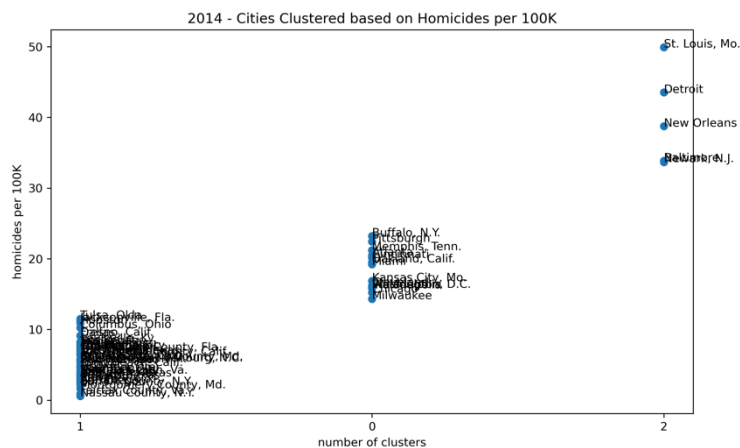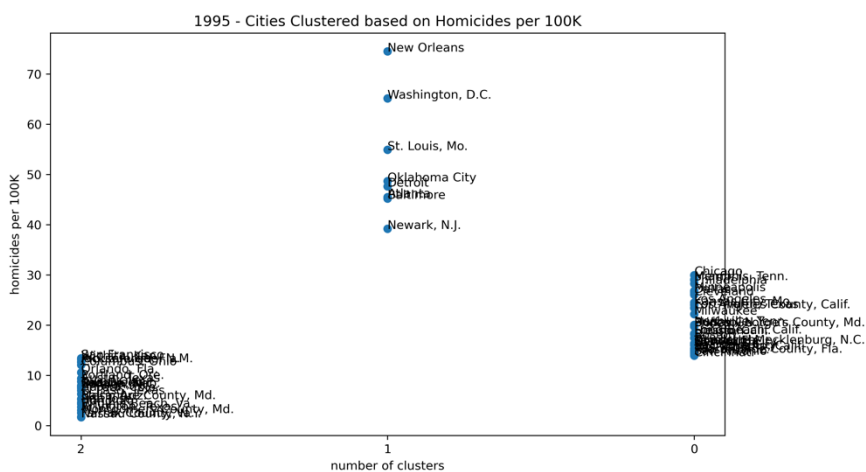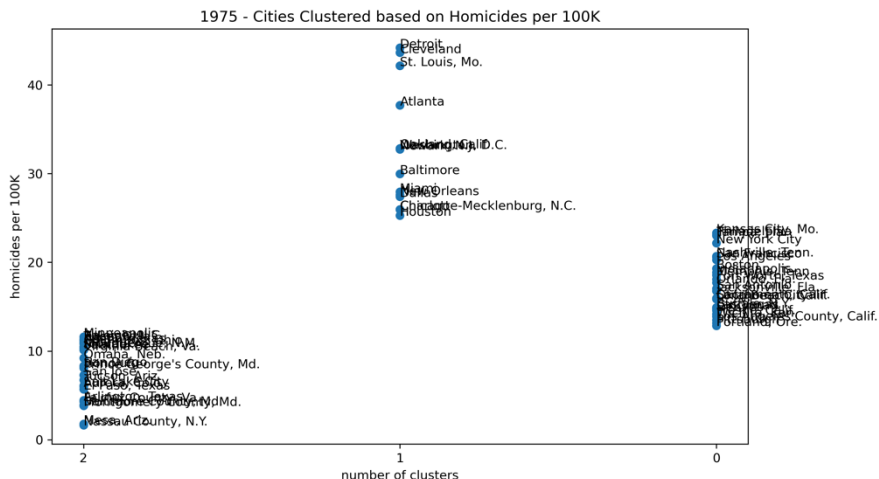

Figure 3 - Geo Heat Map Comparing Homicides per 100k in Major Cities Across the United States in 2014

In 2014 homicide rate varied between 0.57 and 49.91 per 100k. On average the homicide rates were the lowest of the three years indicating a decrease in crime and homicide since 1995.

Overall, the visualizations showed that the homicide rate varied between 1.64 and 44.19 per 100k in 1975, between 1.66 and 74.51 per 100k in 1995 and between 0.57 and 49.91 per 100k in 2014. It was observed that from 1975 to 1995, cities either remained constant (low to low or mid to mid like for example: Los Angeles) or they decreased in homicides per 100k (Cleveland for example). Then from 1995 to 2014, it was observed that most cities, even those that saw an

uptake in homicide rates from 1975 to 1995, shifted and decreased (Salt Lake City for example). Many of the cities have changed from a higher crime group to a lower crime group over the three periods, 1975, 1995, and 2014. While there are certainly outliers that saw an increase, New Orleans for example, the majority did not.

1975 - Cities Clustered based on Homicides per 100K

1995 - Cities Clustered based on Homicides per 100K

2014 - Cities Clustered based on Homicides per 100K

The cluster graphs are a useful tool for understanding changes in crime rates between different groups of cities. They provide a visual representation of the distribution of cities across different crime clusters, allowing for easy comparison between different years.

Discussion

The present study found that crime rates in major U.S. cities have decreased over the past 40 years, with the majority of cities moving to a lower crime cluster from 1975 to 2014. However, it is important to note that the analysis was limited to three specific years (1975, 1995, and 2014) and did not include data on other potential factors that may have influenced the change in crime rates such as policing strategies and demographic changes.

While the year was used as the explanatory variable, it is not clear why the changes in homicide per 100k occurred and it is not necessary to imply that homicide rates will continue to decrease based on year alone. This leads us to question what other variables might be responsible for the shifting of many cities into lower crime clusters over the years. It falls under the domain of knowledge that further research is needed to explore other factors that may have contributed to the decrease in crime rates.

Potential explanations for the decrease in crime rates could include increased police presence in cities, changes in policing strategies such as community policing, and an overall decrease in poverty and unemployment rates. Additionally, demographic changes such as an aging population may also have played a role. However, it is important to note that these are only potential explanations and further research is needed to confirm or disprove these theories.

It is also important to note that there were outliers such as New Orleans that saw an increase in crime rates. Further research could examine why these specific cities did not follow the overall trend of decreasing crime rates. Overall, it is important to consider the limitations of the study when interpreting the results and drawing implications for future research.

Limitations:

The present study has several limitations that should be considered when interpreting the results. These include:

1. Lack of data on policing and demographic information: The dataset used in this study was limited to crime rates and numbers, and did not include information on policing and demographic variables that could have been associated with crime. This limits the ability to make inferences about the data and the potential real-world implications.
2. Limited number of cities and years analyzed: The dataset included data from a limited number of cities, and the analysis was limited to the years 1975, 1995, and 2014. This limits the generalizability of the results.
3. Explanatory variable limitations: The use of year as the explanatory variable may not fully explain the changes in homicide rates, and the trends observed do not necessarily imply that homicide rates will continue to decrease in the future.

4. K-means clustering limitations: The use of K-means clustering has limitations as the clusters are relative and not fixed. This means that the clusters for each year may differ, making it difficult to compare cities across years.
5. Exclusion of data: The analysis excluded data that was not reported over a 12-month period, which may limit the ability to fully examine data from major cities.
6. Time constraint: The analysis was limited to three years instead of 40 years, which may limit the ability to fully examine long-term trends.

These limitations should be taken into consideration when interpreting the results and drawing implications for future research.

Conclusion

Our analysis provides an overview of crime trends in major U.S. cities over the past 40 years. The results indicate a decrease in crime rates in most cities, with the majority of cities moving to a lower crime cluster. However, it is important to note that our data only covers a period of three years (1975, 1995, and 2014) and we excluded data that was not reported over a 12-month period. Additionally, we chose to work with homicides per 100k instead of total homicides as the magnitude of correlation was higher for this measurement. Therefore, further research is necessary to fully understand the underlying causes of these crime trends and to gain a more comprehensive understanding of crime in major U.S. cities. Overall, while our findings suggest a decrease in crime rates in major U.S. cities, it is important to continue to examine and understand the underlying causes of this trend in order to develop effective crime reduction strategies.

Future Research:
The next phase may be to look at more recent years which would be more relevant and maybe do a multiple linear regression to analyze if there are any variables associated with the changes in homicides. Some variables may be policing (overstaffed, understaffed, staffed, not staffed) and demographic details about the cities. Furthermore, answers related questions such as: what factors may be associated with the decline of homicides in the following years 1975, 1995, and 2014? Or what are the trends/changes in homicides in more recent years and what factors may be associated with the trends/changes in homicides in more recent years?