(a) Abstract

The problems that we are targeting is how the crime rate in the neighborhood and the areas affect the real estate prices, analysis on the data that will help us find the root causes of some of the crimes and which could eventually help to find the steps that would rectify those problems.

(b) Introduction

The problem statement can be summarized as "Identifying how crime in the area affects the real estate prizes". If we want to dig a little deeper with this definition then we are also planning to perform root cause analysis on the data and identify the causes of a particular crime in the area and find out what necessary steps can be taken to prevent the crimes. According to us, this problem is vital to solve because there are so many cases in which naive buyers or someone who is unfamiliar with the area ends up buying a house which is not in the safe area. This could be very dangerous as well as prone to a monetary loss. So focusing only on Boston, it is growing very fast and there are so many opportunities in all fields which attract lots of young professionals. Plus, with the high cost of real estate near the city could easily throw someone off to buy a cheaper house in the unsafe area. If there is a system in place for something like this, then it could prevent such issues. This is merely a very simple example, but if steps are not taken to prevent the crime from happening, then this could spread and make the entire place unsafe to be in. There are several datasets available for our purpose. For example, https://data.boston.gov/dataset/crime-incident-reports-august-2015-to-date-source-new-system is the dataset by the government on the crime reports. There are several datasets for property and real estate based on areas too. This could help, fulfill our evaluation needs.

(c) Methodology

1) Dataset

Certain datasets we have analyzed are:

1) Boston Crime Dataset
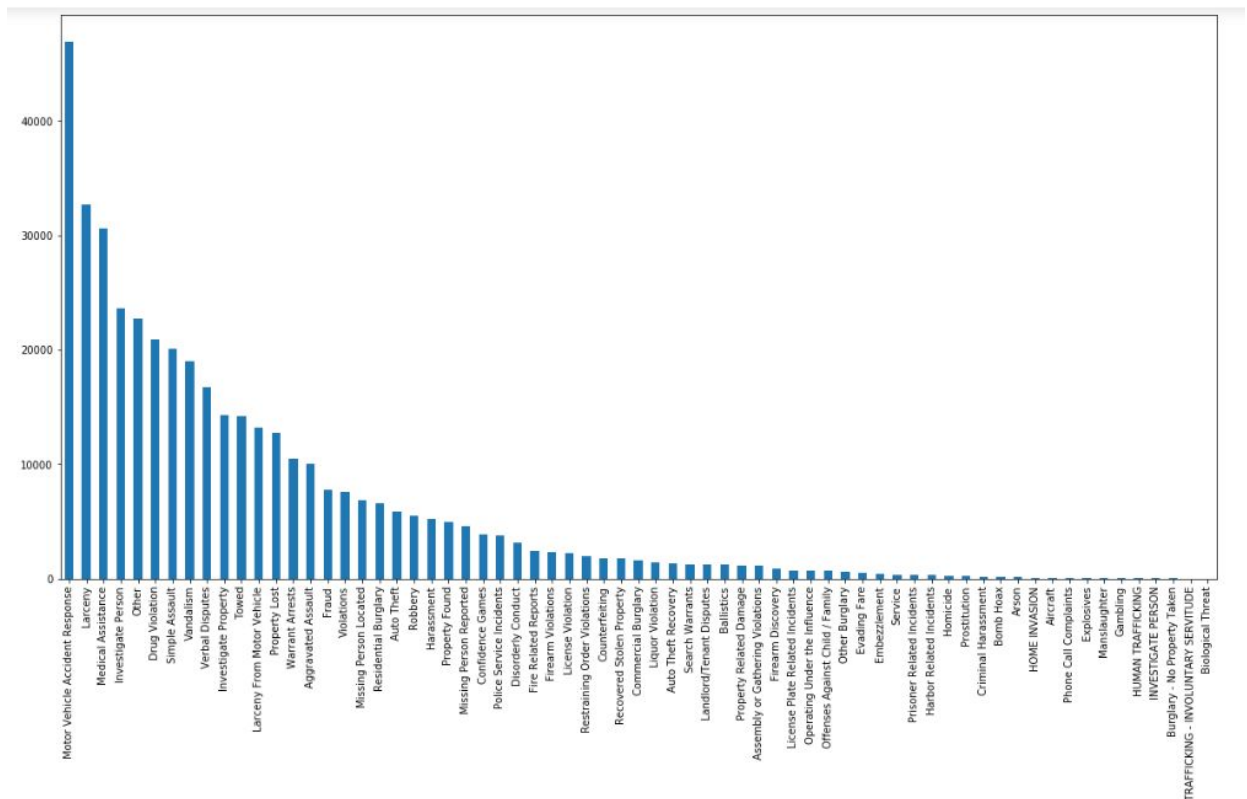   (https://data.boston.gov/dataset/crime-incident-reports-august-2015-to-date-source-new-system)
2) Boston Prices Housing Dataset
   (https://www.kaggle.com/c/boston-housing/overview)

Further, we are researching various datasets from which we can get authentic and good data.
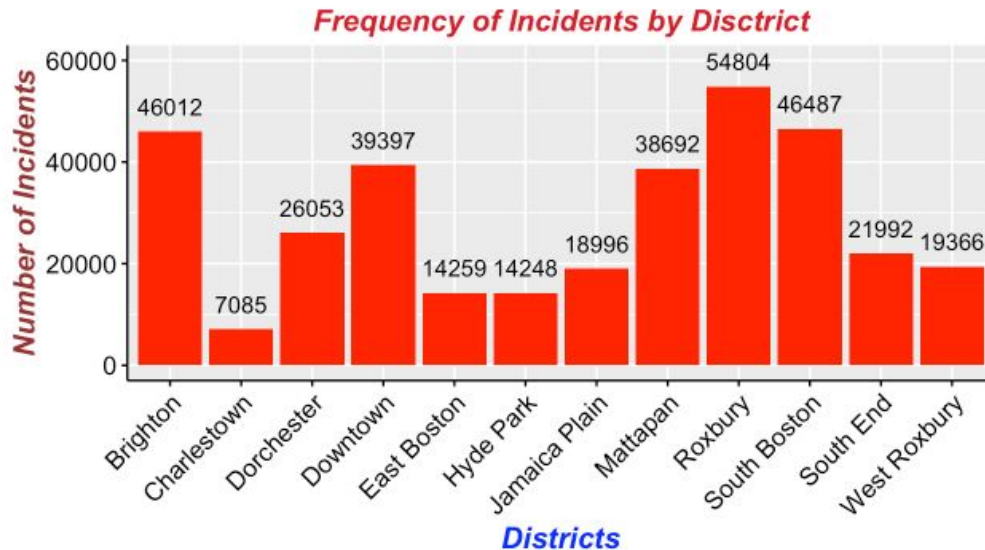
2) Analyzing the dataset:

To start with we are analyzing the crime data and observing what trends we can find from them. For example, below we can see that the most crime that is observed in the Boston city is as shown below:

Also, from the analysis, we found that Washington Street in Boston has the highest crimes registered in the past three years.

Looking at instances of crime by districts:

Frequency of Incidents by District

However, this might not be the entire truth of the safety of the region, there are also impact of types of crimes that need to be analyzed.

Hence, so forth we would like to further analyze the dataset and from those results, we would apply a regression model to see if the crime rates persist how would this affect the housing prices in the future.

3) Predicting the housing rates
Based on the analysis of these two datasets, we would further like to predict the prices of the houses based on the observed data. There might be instances that a region currently has a bit higher crime rate but might decrease in future because of various efforts by the officials and other such factors.

4) Finding the root cause of the crime
Based on the data analysis, we would like to analyze the root cause of the crime. We might refer datasets of other cities for this task and get to know how it can be generalized.

5) Analyzing datasets of other cities and their efforts to decrease the crime rate
By further researching the crime data, there might be instances or areas in other parts of the cities where certain parameters might match and the efforts applied there might also apply here and might bring a reduction in the crime rates.

(d) Code

Code is not in place yet.

(e) Results

Not in place yet.

(f) Discussions

Not in place yet.

(g) Future work

The overall study of crime dataset and the root cause analysis performed on it can be useful for other places. As in, if there are similar types of crimes in other areas as well, and then the steps taken in some areas helped solve those crimes, it could be utilized in the different areas as well by the government to see if there is any difference.

(f) Conclusion

Based on the study performed in this project, it is evident that the crime rates in the area do affect the prices of real estate. So the safer the area, the better the prices of houses are.