Boston Crime Rates in Relation to Hospitals, Police Districts and Property Value

George Gelinas

Introduction

In this project, we attempt to understand crime rates and location within the city of Boston. Using data from the city of Boston that includes police districts, hospitals, property value and crime reports. Within this project, we are focusing on the relationship between average property value and crime rates within police districts. Presuming that property value increases as crime rates decrease. Furthermore, we are exploring hospital locations and computing whether hospital locations are optimal against crime rates. The techniques being used within the project are correlation coefficient and k-means clustering to provide analysis and data to either support these claims or result in new claims.

Datasets

Police Districts This is a database that contains the 12 police districts within the city of Boston. The dataset contains the name of each police district with their respective district number. It also provides addresses, geographical coordinates and zip codes. Hospitals This dataset contains information about hospitals within the city of Boston. It contains hospital names, geographical coordinates, addresses, and zip codes. Property Assessment 2016 This dataset contains tax assessment of property in the city of Boston for the year 2016. Each entry in this dataset contains an address, zip code, owner, mailing address, average land value, average building value, average total value, gross tax, and other descriptors for the property. Crime Reports This data set is from the city of Boston police department that contains incident reports from August 2015 to todays date. For every entry in this data set contains a incident number, offense code, offense description, district, date, street address and geographical coordinates. The data set collected from the Crime Reports are from August 2015 until 18 November 2016. Each dataset that was collected and used for this project were collected publicly from the city of Boston.

Correlation of Property Value and Crime

In order to determine a relationship between property value and crime rates we computed the relationship using correlation coefficient. We calculated average property value for each respective police district by clustering zip codes within those district zones. Furthermore, we also grouped the crimes within the crime incident report to their respective district. We calculated the correlation coefficient as -0.3138711 and our p-value as 1.0.

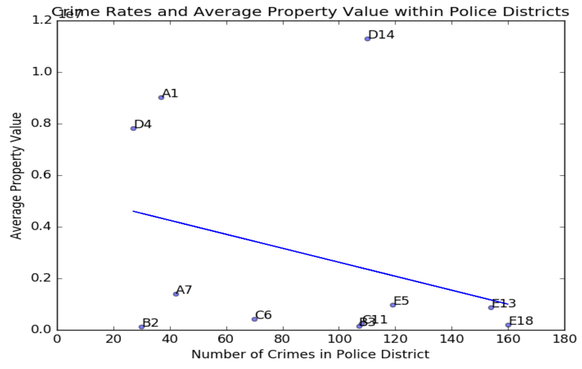

Figure 1: Visual of the correlation between Average Property Value and Crime Rates
Average Property Value (1^e7): 0.2 = 2,000,000 Million

With these results it shows that based on our calculations there is a moderate negative correlation between Average Property Value and Crime Rates with a high p-value meaning that this observation is a non-significant result. This result suggests that the null hypothesis may be true and that as crime rates increase property value decreases.

K-means with Hospital Locations

For this analysis, we used crime incidents and hospital geographical coordinates from their respective dataset. Using the k-means algorithm divides the dataset of crime incidents into k clusters which is set to the number of hospitals within the hospital dataset.

These k-clusters optimize some measurement, in this case the measurement is the distance between existing hospitals and the k means. When the algorithm finishes, we will have a set of optimal hospital locations in regards to location of crime/incidents.

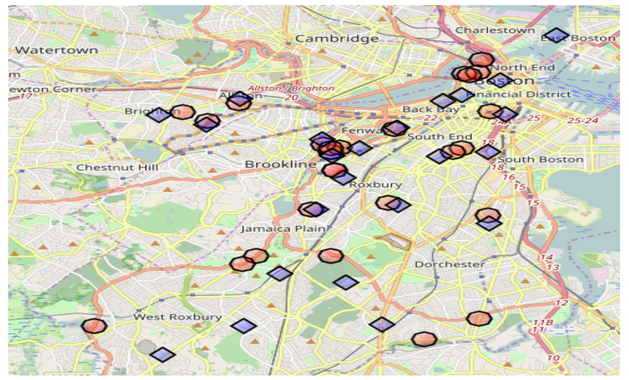

Figure 2: A map containing current hospital locations and optimal hospital locations. Blue squares represent the optimal hospital locations & Red circles represent the current hospital locations

It seems that current hospital locations are not far off from the optimal locations based on crime locations. From this data, it shows that hospitals are in a good location in regards to crime. This analysis raises other questions such as “Do city planners place hospitals near crime in order to shorten distances and time from the location of the incident to the nearest hospital?”

Conclusion

Using the data from the city of Boston, findings show that crime location and frequency may have relationships with property value and hospital locations. Although the correlation is not strong, future work would be to include more zip code regions to see if there is a difference in correlation with multiple areas tested instead of grouping property value within 12 districts. In regards to hospital locations it is interesting to discover that hospitals are closely located to crime. Future work could consider calculating the correlation between hospital and crime locations and determine whether there is a relationship or not.

Instructor: Andrei Lapets