PLACE TO SETTLE IN BOSTON

Sashank Kishore

Capstone Project
IBM Data Science Professional Certificate

1. Business problem –

To find good location to reside in Boston, Massachusetts with low crime rate and good Asian restaurants.

2. Introduction –

Boston is the capital and most populous city of the Commonwealth of Massachusetts in the United States, and the 21st most populous city in the United States. The city proper covers 49 square miles (127 km2) with an estimated population of 692,600 in 2019.

Boston is sometimes called a "city of neighborhoods" because of the profusion of diverse subsections; the city government's Office of Neighborhood Services has officially designated 23 neighborhoods.

Present day Boston is hub for population of various religion and ethnicities.

Even though, unlike many major American cities, Boston has seen a great reduction in violent crime since the early 1990s, there are still many cases registered daily. Boston's low crime rate since the 1990s has been credited to the Boston Police Department's collaboration with neighborhood groups. This helped lead in part to what has been touted as the "Boston Miracle". Murders in the city dropped from 152 in 1990 (for a murder rate of 26.5 per 100,000 people) to just 31—not one of them a juvenile—in 1999 (for a murder rate of 5.26 per 100,000).

For a new person or family to start their life in Boston, they need to find a place with low crime rate. Accompanying with low crime rate, the location should be filled with different facilities. One such important facility is the availability of restaurants. Boston is filled with restaurants of different cuisines such as Italian, Indian, American and many more.

3. Data –

3.1 Data Scraping and Cleaning

Source 1 –

https://data.boston.gov/dataset/crime-incident-reports-august-2015-to-datesource-new-system

Crime incident reports are provided by Boston Police Department (BPD) to document the initial details surrounding an incident to which BPD officers respond. This is a dataset containing records from the new crime incident report system, which includes a reduced set of fields focused on capturing the type of incident as well as when and where it occurred. Records in the new system begin in June of 2015.

The CSV contains 17 Columns and 494021 Rows.

Data Types –

COLUMN	DATA TYPE
INCIDENT_NUMBER	Int
OFFENSE_CODE	Int
OFFENSE_CODE_GROUP	String
OFFENSE_DESCRIPTION	String
DISTRICT	String
REPORTING_AREA	String
SHOOTING	String
OCCURRED_ON_DATE	Date
YEAR	Int
MONTH	Int
DAY_OF_WEEK	String
HOUR	Int
UCR_PART	String
STREET	String
Lat	Float
Long	Float
Location	String

Source 2 –

https://en.wikipedia.org/wiki/Boston_Police_Department

Contains the details regarding the different district codes. Created a table from the above code to get the district names.

Here, crime details were given based on the district code of different police stations and not on district zip code. So in rare cases where there was only one police station for few districts, we grouped them together.

With the present data we have extensive records of crime which can be analyzed with the help of K Mean Method and clustered.

Source 3 – Foursquare API

Foursquare API was used to find the nearby destinations.

3.2 Data Preparation

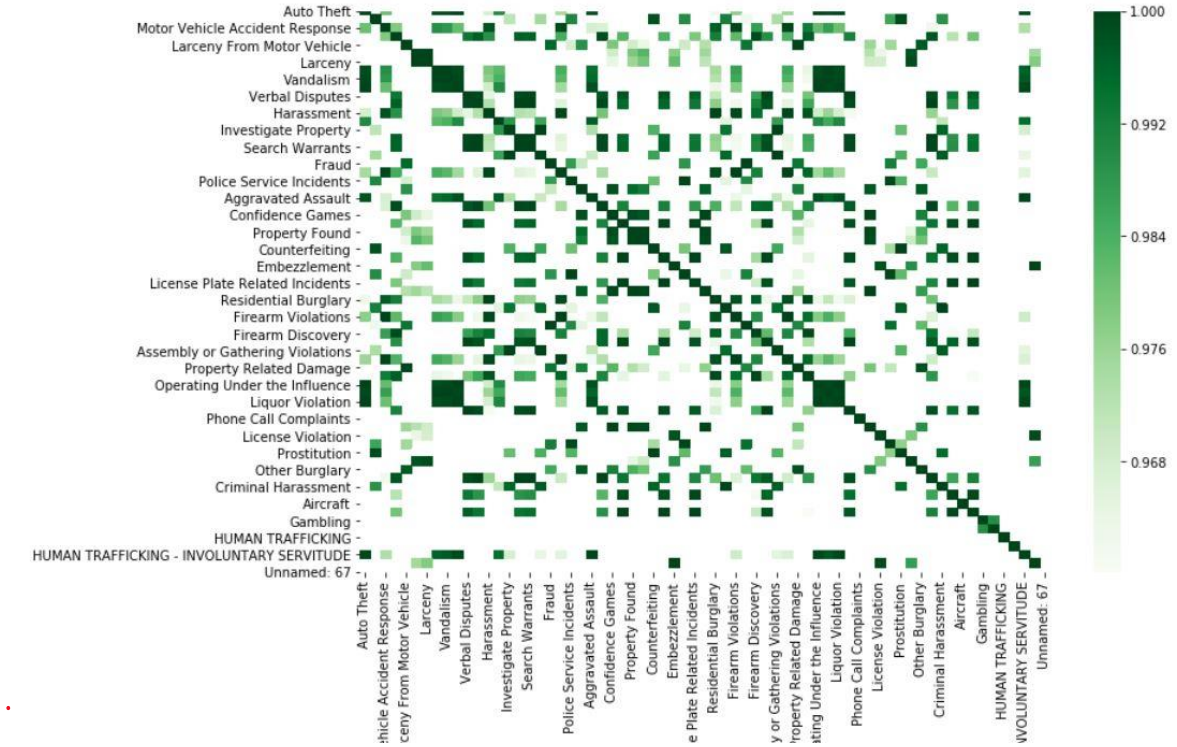
Unwanted columns were removed from the Boston crime data. Rows with no data was removed.

After cleaning the data from first source, it was merged with the second table to get the district name from the district code.

4. Methodology – K Mean Method

After preparing the data, K Mean Method was used to cluster the data. Using the elbow method.

Here we can even find the correlation between different crimes. This can shed come light on cause and effect of different crimes.



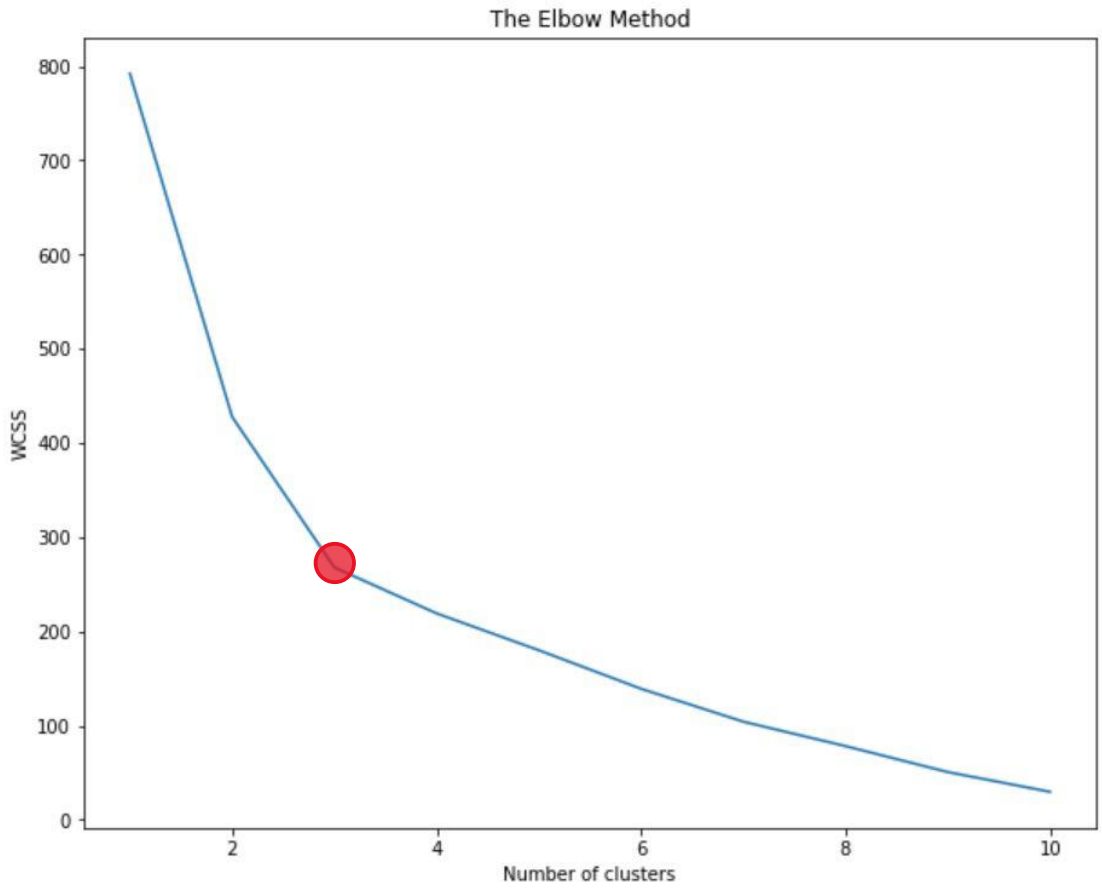
The above diagram show the correlation between various crimes.

Through observation, cluster 1 districts have low crime rate.
Using foursquare api, we were able to find the required destination with in the cluster 1 districts.
Then the data obtained was cleaned based on required filters like specific type of restaurant like ‘Indian’ or based on regions like ‘asian’.

5. Results –

KMM –

Using the elbow method, we found the optimal number of clusters to be 3.



Cluster 1 –

Location	Auto Theft	Simple Assault	Motor Vehicle Accident Response	Violations	Larceny From Motor Vehicle	Towed	Larceny	Ir
Hyde Park	285	905	2694	373	661	411	1264	
Allston,Brighton	344	1101	3213	313	810	1883	2118	
Beacon Hill	121	356	1079	110	369	442	647	
East Boston	224	853	1780	363	479	737	881	
West Roxbury,Roslindale	219	684	2214	288	574	392	1022	
South Boston	530	1391	3121	241	1159	1725	2411	
Jamaica Plain	341	950	2500	470	1033	1063	1685	

7 rows × 68 columns

Cluster 2 –

Location	Auto Theft	Simple Assault	Motor Vehicle Accident Response	Violations	Larceny From Motor Vehicle	Towed	Larcen
Mattapan,North Dorchester	711	2347	4387	993	1020	836	167
Roxbury,Mission Hill	970	3286	6748	1516	1778	1208	368
Dorchester	918	2756	6326	1154	1551	1532	299

3 rows × 68 columns

Cluster 3 –

Location	Auto Theft	Simple Assault	Motor Vehicle Accident Response	Violations	Larceny From Motor Vehicle	Towed	Larceny
Back Bay,South End,Fenway	714	2646	3728	469	2248	2412	9200
Downtown	411	2823	3053	318	1469	1557	6077

2 rows × 68 columns

Foursquare

Going through all the districts in cluster 1, we found that **Allston, Boston,** has most asian restaurents.

6. Conclusion –

Through this analysis were able to find 7 districts in Boston which are much safer than the rest of the city. This was achieved with the help of KMM. To further find our interests, we used Foursquare API to find various destination in the required district. Finally a person who wants to settle in a low crime area and area filled with Asian restaurants would choose **Allston, Boston** as their residencial location.

This analysis was limited by the aount of data that can be retrived through Fousquare API(personal account).

7. Bibliography –

- Boston Government Dataset - https://data.boston.gov/dataset/crime-incident-reports-august-2015-to-datesource-new-system
- Wikipedia - https://en.wikipedia.org/wiki/Boston_Police_Department
- Foursquare API - https://foursquare.com/developers/