

Capstone Project – Battle of the Neighborhoods

Analysis of the Neighborhoods in Boston, MA

Abby Antoo Antony

INTRODUCTION

Safety and access to necessities are the biggest concerns when thinking about moving to a new location.

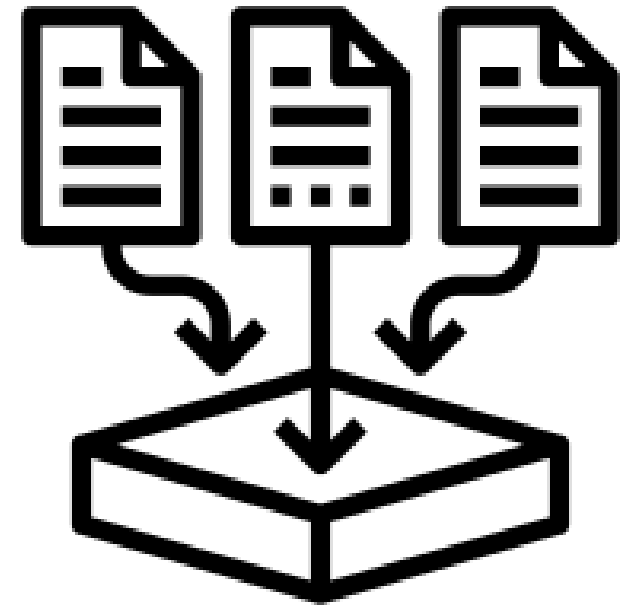
Goals of this project/analysis:

- Identify the safest neighborhood
- Cluster venues per street using K-means



DATA ACQUISITION

The data is extracted from [Analyze Boston](#), which is an open data hub of the city of Boston. The data set that will be used is [Crime Incidents Report](#). These are reports provided by the Boston Police Department to document the initial details surrounding an incident to which BPD officers respond.





Cleaning the Data

The cleaning process is split into 5 parts. They are listed as follows:

1. Selecting the latest year for analysis i.e 2019
2. Adding the column 'District_Names' to identify the names of the different neighborhoods
3. Dropping columns that are not needed for the analysis
4. Dropping missing values.
5. Converting the dataframe to a pivot table and adding the total column.

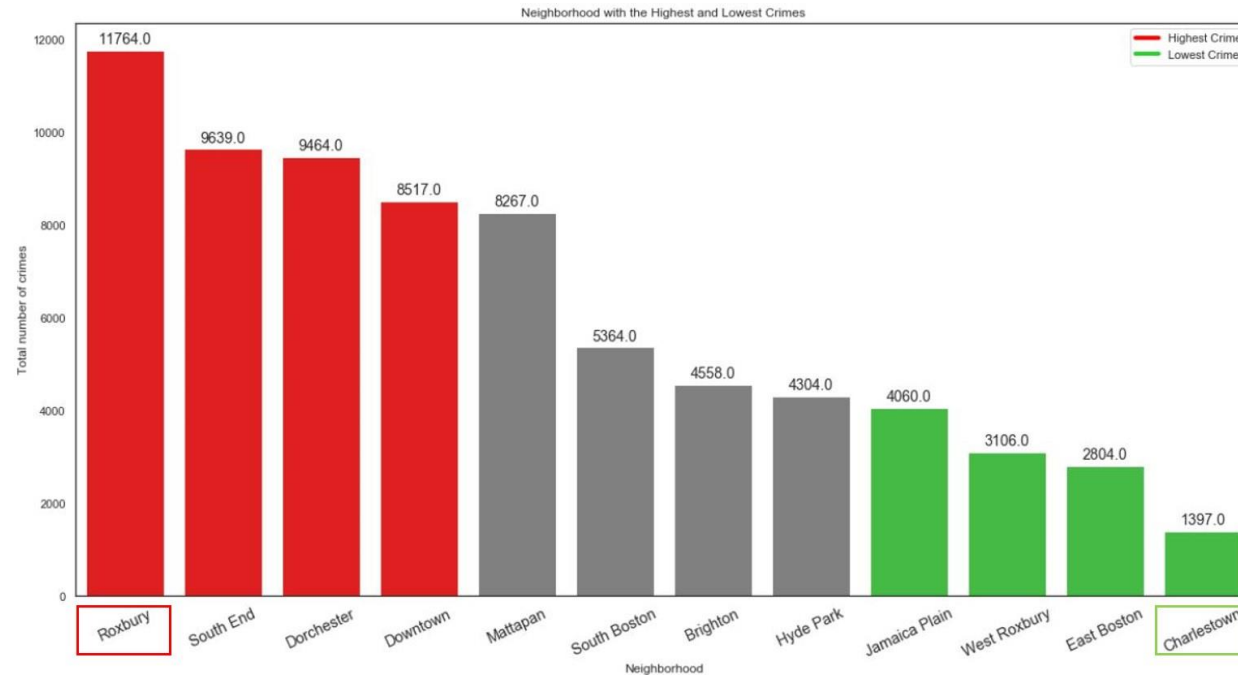
Methodology

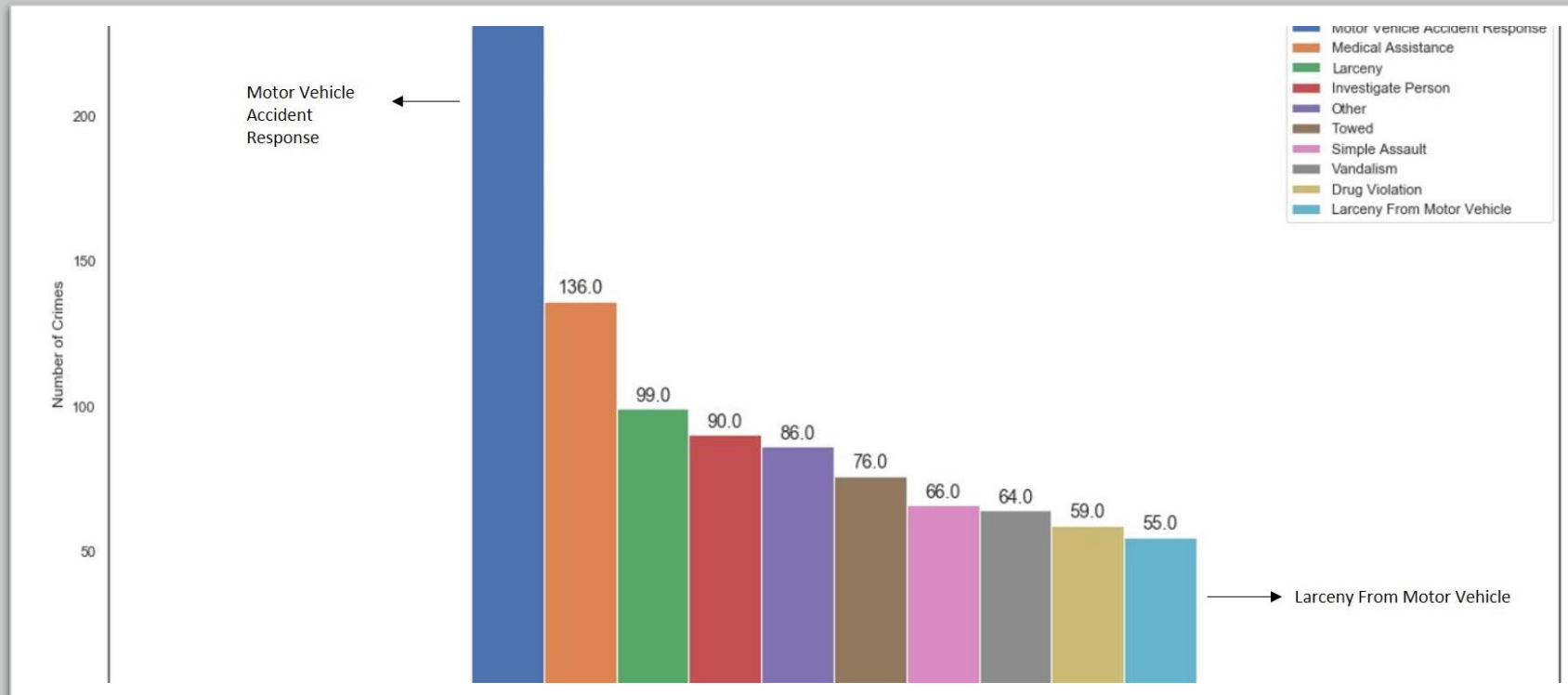
Exploratory Analysis

Visualizing neighborhoods with the highest and lowest reported number of crimes.

***Highest at 11764 – Roxbury**

***Lowest at 1397 - Charlestown**

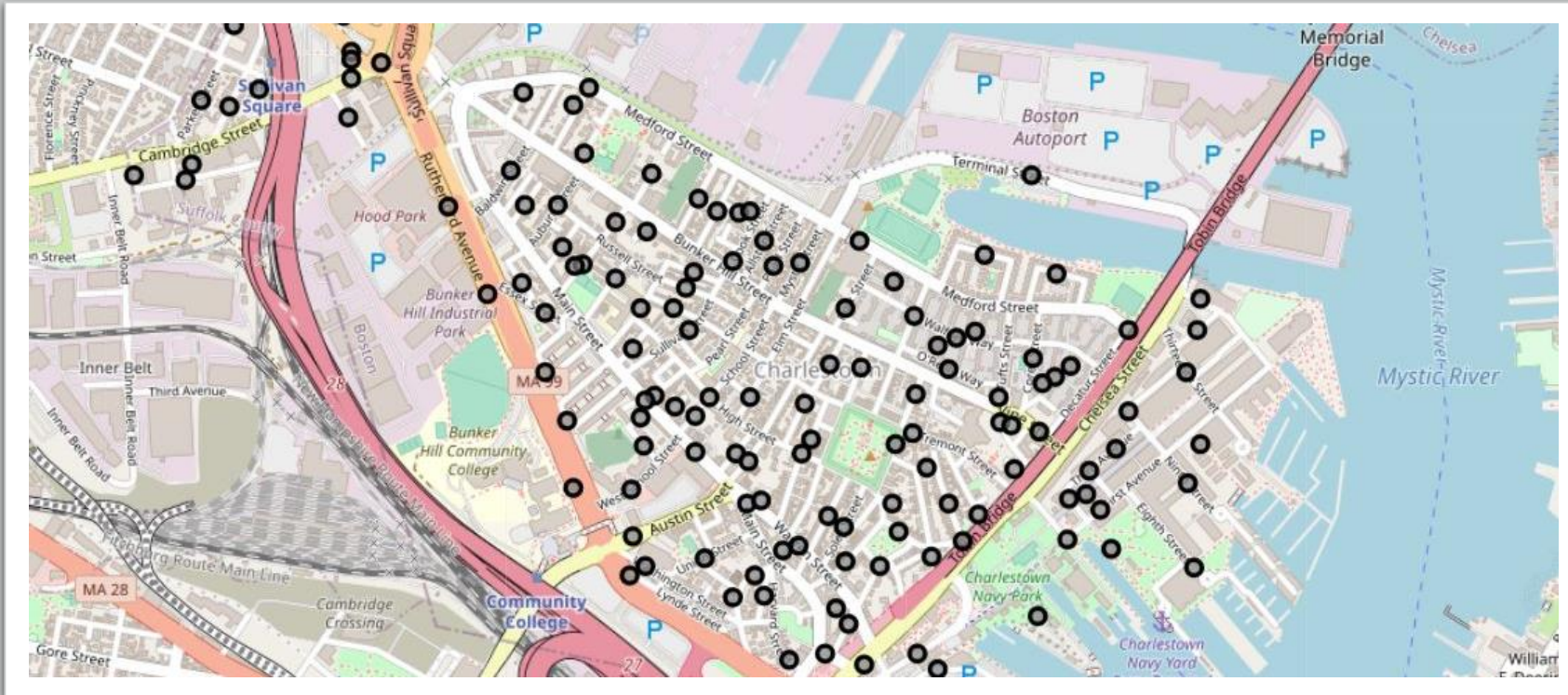




Number of crimes in Charlestown based on the Offense Category.

***Highest at 232 – Motor Vehicle Accident Response**

***Lowest at 55 – Larceny From Motor Vehicle**



Mapping Charlestown using Folium \Rightarrow Charlestown has 138 unique streets which can be seen on the map above.

K-means - Clustering

- Connect to Foursquare API to find venues within 500-meter radius of each street in Charlestown. The data extracted is in the form of JSON.
- JSON is converted into a dataframe and assigned to variable Charlestown_data.
- The data is One Hot encoded to convert categorical variables into binary form.
- Run K-means – an algorithm that segregates data and clusters them based on similar traits given that the cluster sizes are predefined. ($k = 5$)



Results

1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
Pizza Place	Bus Station	Pool	Convenience Store	Theater	Gym / Fitness Center	Gym	Music Venue	Café	Yoga Studio
Convenience Store	Pizza Place	Deli / Bodega	Tennis Court	Pool	Café	Gastropub	Gym	Music Venue	Athletics & Sports
Convenience Store	Pizza Place	Deli / Bodega	National Park	Café	Music Venue	Gastropub	Monument / Landmark	Pool	Gym
Convenience Store	Pizza Place	Boat or Ferry	National Park	Music Venue	Candy Store	Café	Gastropub	Monument / Landmark	Donut Shop
Convenience Store	Deli / Bodega	Martial Arts Dojo	Donut Shop	Pizza Place	Candy Store	Café	National Park	Music Venue	Gym
Convenience Store	Deli / Bodega	Boat or Ferry	Donut Shop	Pizza Place	Park	National Park	Café	Music Venue	Gastropub

Cluster 1

- 18 out of 138 streets are in this
- Current or potential residents have access to pizza places, deli shops and coffee shops. Other venues include music venues, tennis courts and gyms but they are not as common.
- Pizza places appears 14 times in the 1st and 2nd most common venues.

	Street	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
2	LOWNEY WAY	History Museum	Gastropub	Pub	Donut Shop	Park	National Park	Grocery Store	Convenience Store	Coffee Shop	Seafood Restaurant
5	EDEN ST	Pizza Place	Yoga Studio	Gym	Pharmacy	Pet Store	Coffee Shop	Donut Shop	Café	Shopping Mall	Gastropub
6	THOMPSON SQ	Pizza Place	Café	Pub	Coffee Shop	Gastropub	Pet Store	Plaza	Park	History Museum	National Park
11	SALEM ST	Pub	Convenience Store	Coffee Shop	Yoga Studio	Gastropub	Donut Shop	Pharmacy	Deli / Bodega	Pet Store	National Park
14	CHARLES RIVER AVE	Pub	Pizza Place	Donut Shop	Park	Brewery	Moroccan Restaurant	Playground	Plaza	Restaurant	Lounge
18	AUSTIN ST	Coffee Shop	Pizza Place	Pub	Yoga Studio	Shopping Mall	American Restaurant	Bank	Convenience Store	Donut Shop	Gastropub

Cluster 2

- Largest cluster – Consists of 60 out of 138 streets.
- Current or potential residents have access to pubs, pizza paces, convenience stores, café's and pharmacies.
- Pubs appear 37 times in the 1st , 2nd , and 3rd most common venue columns.

	Street	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
4	MONUMENT ST	Boat or Ferry	Surf Spot	Pizza Place	Donut Shop	National Park	Music Venue	Candy Store	Café	Monument / Landmark	Gastropub
7	BUNKER HILL ST	Boat or Ferry	Donut Shop	History Museum	National Park	Monument / Landmark	Playground	Convenience Store	Park	Candy Store	Café
8	DECATUR ST	Park	Surf Spot	Grocery Store	Donut Shop	Tourist Information Center	Martial Arts Dojo	Convenience Store	Gym Pool	Gastropub	Boat or Ferry
10	VINE ST	National Park	Café	History Museum	Boat or Ferry	Donut Shop	Park	Gastropub	Surf Spot	Grocery Store	Monument / Landmark
12	WALFORD WAY	Donut Shop	Café	Boat or Ferry	Gym	Convenience Store	National Park	Candy Store	Gastropub	Monument / Landmark	Grocery Store
16	EIGHTH ST	Café	History Museum	Harbor / Marina	Grocery Store	Park	Seafood Restaurant	Surf Spot	Boat or Ferry	Beer Garden	Convenience Store

Cluster 3

- Second largest cluster – Consists of 34 out 138 streets.
- Current or potential residents have access to history museums, national parks, boats & ferries and donut shops.
- The other venues that are not very common are pizza places, gyms and grocery stores to name a few.

	Street	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
1	AUBURN ST	Gym / Fitness Center	Bus Station	Pool	Theater	Pizza Place	Gym	Café	Park	Yoga Studio	Dive Bar
9	RUTHERFORD AVE	Bus Station	Gym / Fitness Center	Recording Studio	Food	Bus Stop	Pool	Park	Gym	Café	Bank
13	OAK ST	Gastropub	Bus Station	Pool	Theater	Pizza Place	Gym / Fitness Center	Gym	Café	Park	Dive Bar
25	CAMBRIDGE ST	Bus Station	Gym	Gym / Fitness Center	Donut Shop	Liquor Store	Metro Station	Coffee Shop	Café	Bus Stop	Park
35	BUNKER HILL INDUSTRIAL PA	Bus Station	Pool	Theater	Pizza Place	Metro Station	Food	Gym	Park	Gastropub	Yoga Studio

Cluster 4

- Consists of 25 out of 138 streets.
- Current or potential residents have access to gyms, public transportation, donut shops, café's and liquor stores to name a few.

	Street	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
146	ALFORD ST	Taxi	Restaurant	Yoga Studio	Grocery Store	Diner	Discount Store	Dive Bar	Dog Run	Donut Shop	Electronics Store

Cluster 5

- Consists of only 1 out of the 138 streets.
- This cluster is an outlier.
- Common venues include – Taxi stations, restaurant, yoga studio and grocery stores to name a few.

Discussions

- This analysis aims to help any individual or group of individuals who are planning on moving into the city of Boston.
- The first goal of the analysis was to find the safest neighborhood – Charlestown
- The second goal was to cluster different venues per street in the neighborhood having similar traits.
- Cluster 1 & 2 - For individuals who like to eat out a lot and have fun with access to convenient stores.
- Cluster 3 – For individuals who enjoy hikes, history, boats, ferries and a bit of surfing.
- Cluster 4 – For individuals who enjoy fitness as well as use public transport as their primary source of transportation.



Conclusions

Moving into a new city or country has so many factors to it. This analysis is an indicator of how technology can be utilized to gain additional insights on the city or country one is planning on moving to. Couple lines of code narrowed down the safest neighborhood as well the venues around that neighborhood making the whole process of moving easier for an individual. They could use their time to look into the other factors that were not taken into consideration like the cost of living and housing, which deserves their own analysis. But as a start, anyone who is planning to move to the city of Boston, Charlestown is the place to be.

