Capstone Project - The Battle of Neighborhoods: Boston's Future

SUBMITTED BY: DARKO KNEZEVIC

Introduction

- Boston is the capital of the U.S. state of Massachusetts.

- Key facts:
  - A growing economy and a great place to acquire an education.
  - Some of the most prestigious schools within the U.S.A are located in Boston.
  - A lot of large companies are seeking to place their firms in Boston.
  - A global city that is placed among the top 30 most economically powerful cities in the world.
  - Boston's colleges and universities exert a significant impact on the regional economy.
  - Attracts hundreds of thousands of students from around the world, who contribute billions of dollars annually to the city's economy.
  - Highly innovative for a variety of reasons, including the presence of academia, access to venture capital, and the presence of many high-tech companies. Tourism also composes a large part of Boston's economy.

Problem Statement

- The amount of crime needs to be understood for Boston to continue to thrive as a major hub for innovation, tourism, academia, and employment.
  - If the Boston Police Department knew more about which areas to concentrate on, then Boston's citizens, tourists, and students would feel much safer.
  - Certain areas within the city could be targeted for more protection by the Boston Police Department based on data.

Objective

- The project aims to select the safest and least safest areas in Boston based on the total crimes, explore the venues in these areas and finally cluster the areas using k-mean clustering.

- The report will be targeted to the Boston Police Department who are looking to analyze crime incidents, as the city continues to be a leader in so many fields.
  - The crime statistics will provide an insight for the Boston Police Department towards their approach in dealing with different areas of city.
  - The most common venues in the safest areas will be explored.

Data

- Based on definition of our problem, factors that will influence our decision are:
  - The total number of crimes committed in each major area during the most recent year available. For example, acquiring data from the most recent year available will be done to do analysis.
  - The most common venues in the area that is the safest.
  - The total amount of crime that have been committed from the year 2015 to 2018 will be analyzed for clustering purposes. More data will allow for more data analysis to be done.

- Following data sources will be needed to extract/ generate the required information:
  - Preprocessing a real-world data set from Kaggle showing the Boston Crimes from the last few years. Kaggle, a subsidiary of Google LLC, is an online community of data scientists and machine learning practitioners. Kaggle allows users to find and publish data sets, explore and build models in a web-based data-science environment, work with other data scientists and machine learning engineers, and enter competitions to solve data science challenges.

Data Continued

▪The Boston crime file provides information about the following:

INCIDENT_NUMBER	YEAR	LOCATION
OFFENSE_CODE	MONTH	
OFFENSE_CODE_GROUP	DAY_OF_WEEK	
OFFENSE_DESCRIPTION	HOUR	
DISTRICT	UCR_PART	
REPORTING_AREA	STREET	
SHOOTING	LATITUDE	
OCCURRED_ON_DATE	LONGITUDE	

▪Data Set URL: https://www.kaggle.com/AnalyzeBoston/crimes-in-boston
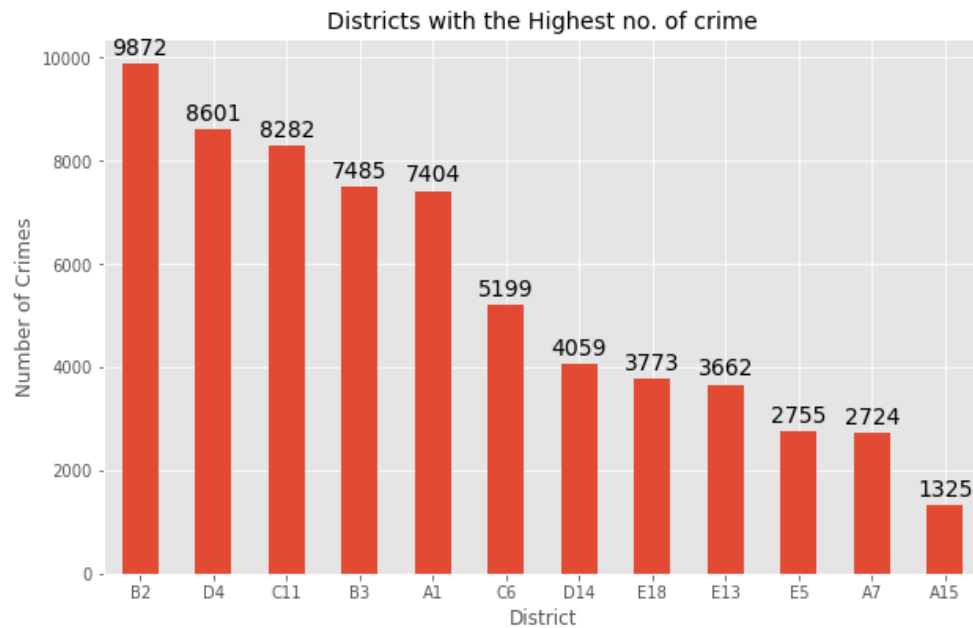
Methodology

- Data Cleansing:
  - From the Boston crime data, the data from the year 2018 was initially analyzed to understand key drivers in the most recent year. However, as the study progressed in order to effectively do k-means clustering more data from previous years was required.
  - All the data values that didn't have a district associated with an incident were removed from the dataset for certain types of analysis.
  - Missing data were coming from reporting area, latitude and longitude.

Statistical Summary of Crime in 2018

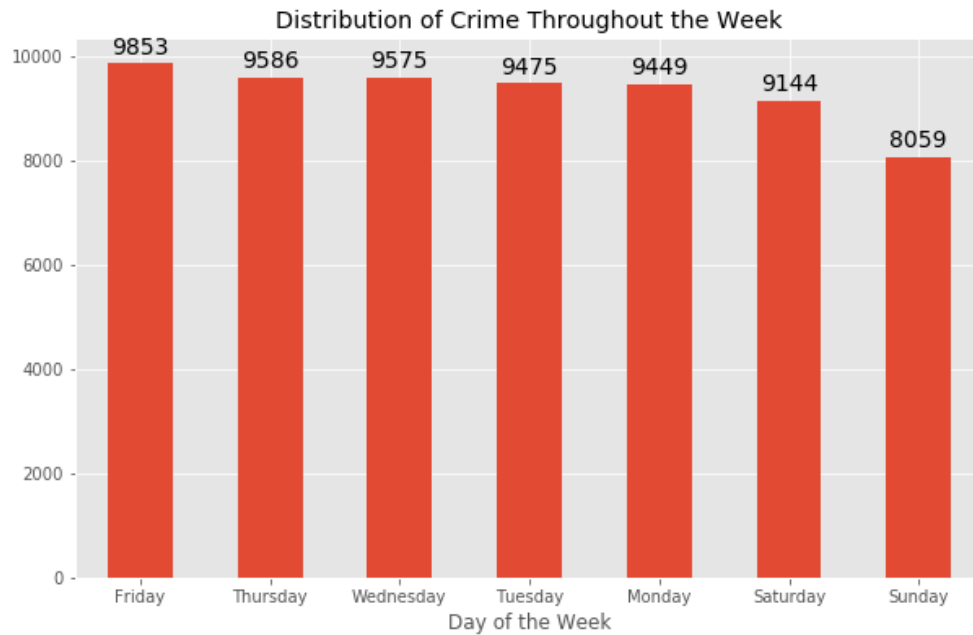|	OFFENSE_CODE	YEAR	MONTH	HOUR	Lat	Long
count	65141.000000	65141.0	65141.000000	65141.000000	60968.000000	60968.000000
mean	2357.715525	2018.0	4.662747	13.099384	42.301265	-71.049630
std	1174.774400	0.0	2.297052	6.261562	0.945161	1.528421
min	111.000000	2018.0	1.000000	0.000000	-1.000000	-71.178674
25%	1102.000000	2018.0	3.000000	9.000000	42.296907	-71.097387
50%	3006.000000	2018.0	5.000000	14.000000	42.325321	-71.077783
75%	3205.000000	2018.0	7.000000	18.000000	42.348406	-71.062563
max	3831.000000	2018.0	9.000000	23.000000	42.395042	-1.000000

Approximately 65,141 crime were identified in the City of Boston that have been assigned to a police district in 20180
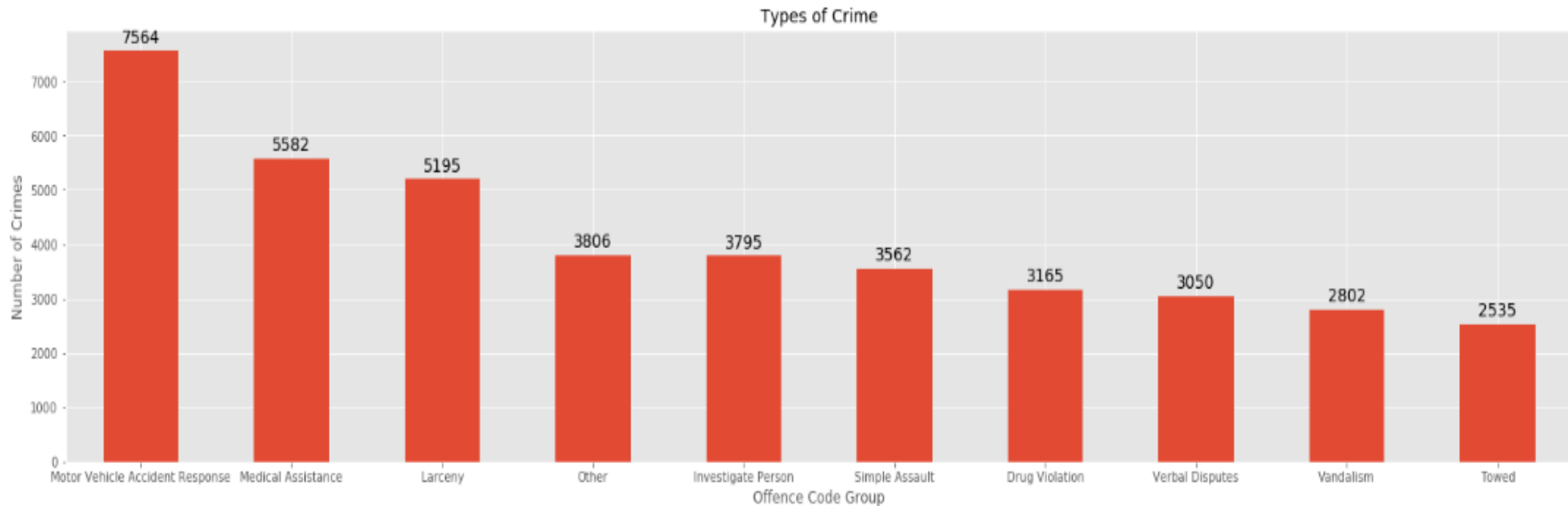
Exploratory of Data

Districts with the Highest no. of crime

- The top five districts were the most crime occurs is at B2, D4, C11, B3, A1 in 2018.

- Most of the crime that is occurring is located in the middle of the Current Boston Police Districts Map.

Exploratory of Data Continued
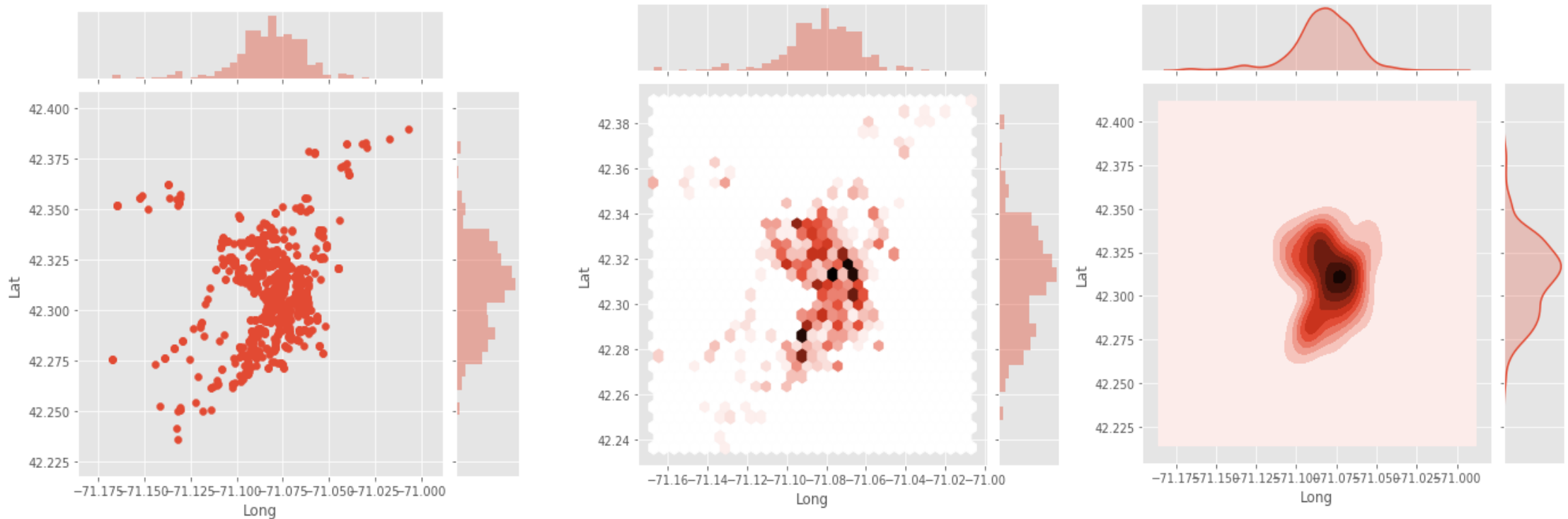

Distribution of Crime Throughout the Week

- More crime occurs as the week goes on from Monday to Friday with a reduction in the amount of crime during the week in 2018.

Exploratory of Data Continued

A large spike in the amount of crimes have been related to motor vehicle accident response, medical assistance, and latency in 2018.

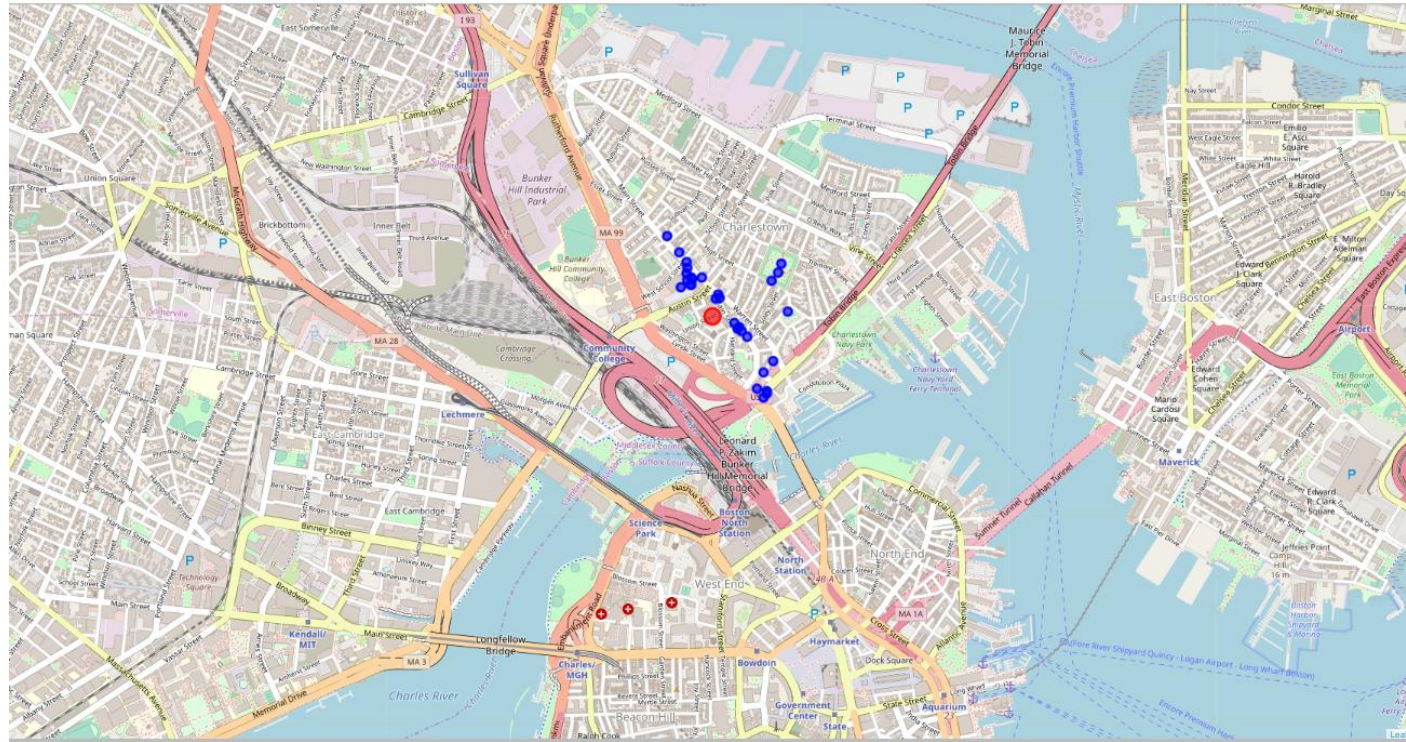# Exploratory of Data Continued

Most of the shooting in Boston from 2015 to 2018 have been located in the centre of the Boston Police Districts.
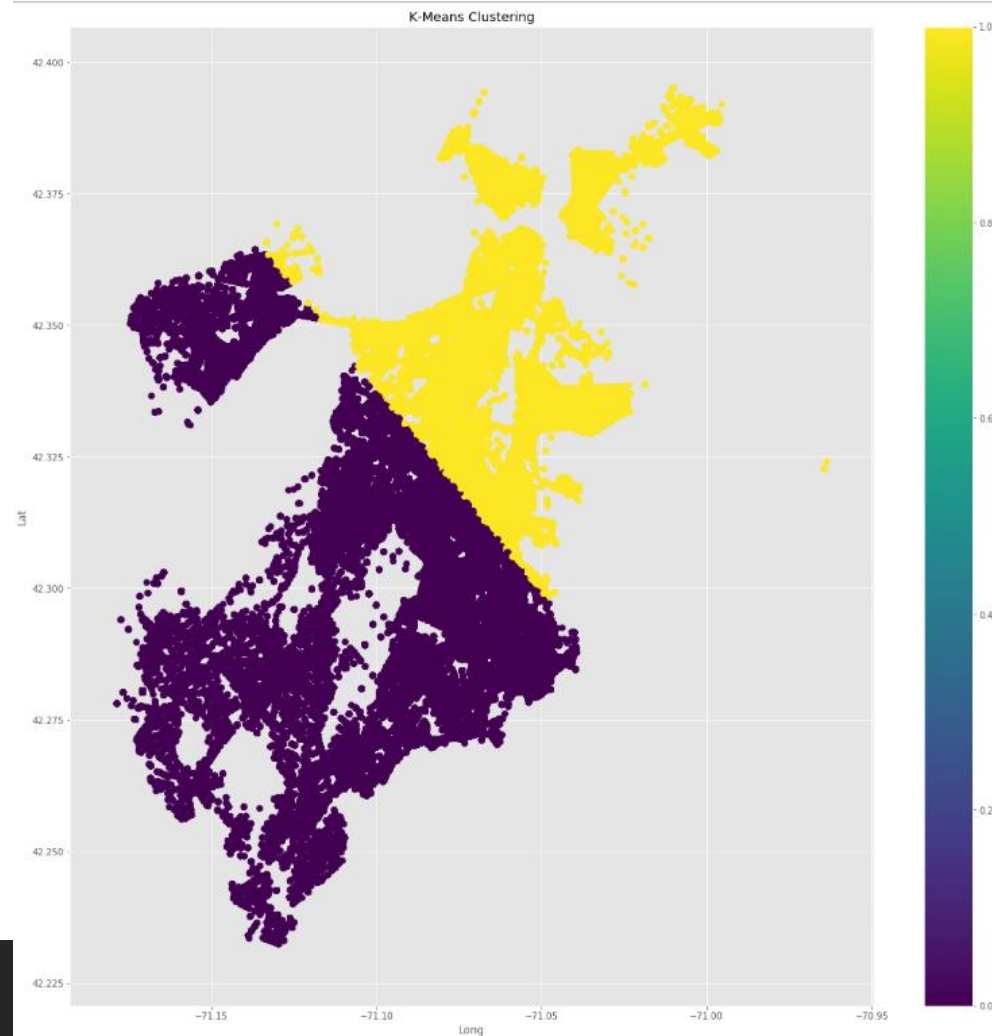
Common Venues in the Safest Area – A15 (Charlestown)

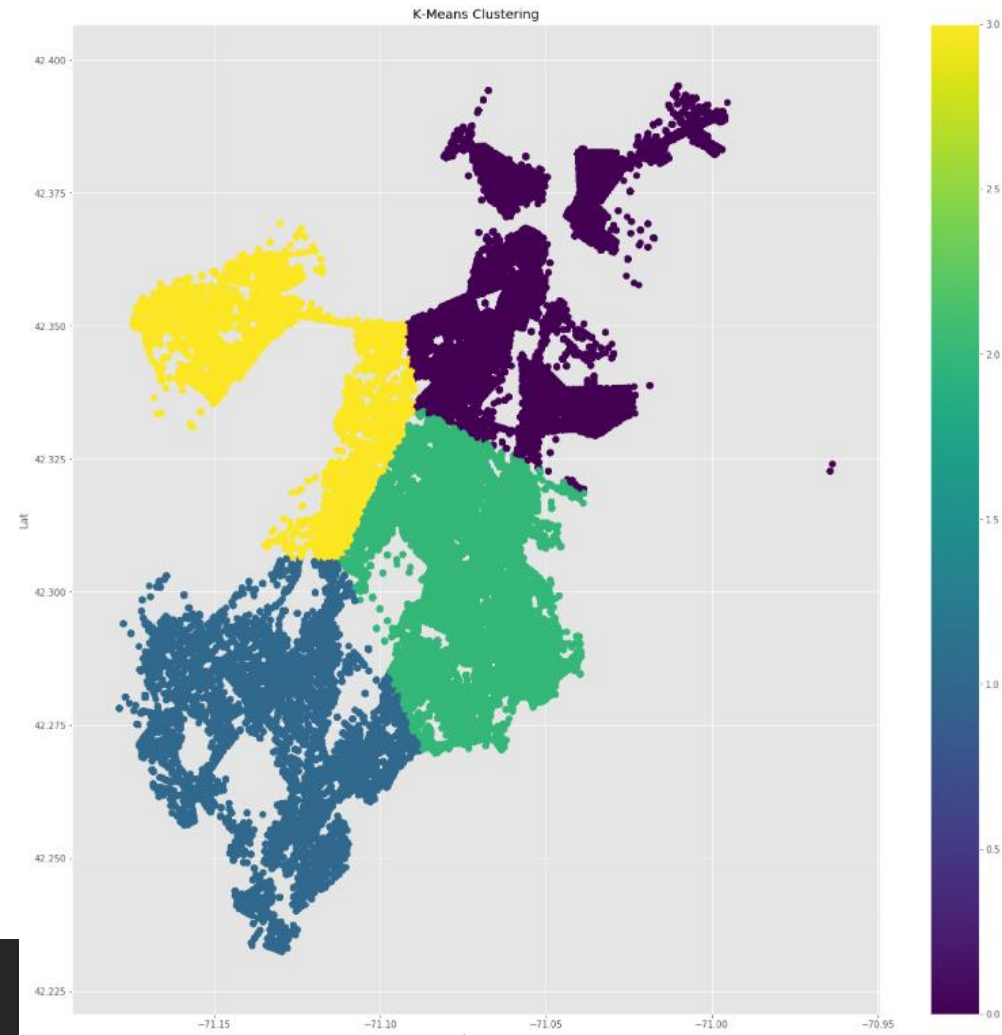|	name	categories	address	crossStreet	lat	lng	labeledLatLngs	distance	postalCode	cc	city	state	country	formattedAddress	id
0	Warren Tavern	Pub	2 Pleasant St	at Main St	42.374108	-71.063143	[{'label': 'display', 'lat': 42.374108, 'lng':...	98	02129	US	Boston	MA	United States	[2 Pleasant St (at Main St), Boston, MA 02129,...	44394daff964a5201c321fe3
1	Figs	Pizza Place	67 Main St	at Winthrop St.	42.373588	-71.062463	[{'label': 'display', 'lat': 42.37358769797685...	173	02129	US	Charlestown	MA	United States	[67 Main St (at Winthrop St.), Charlestown, MA...	4b11c019f964a520d08323e3
2	Whole Foods Market	Grocery Store	51 Austin Street	NaN	42.375511	-71.065976	[{'label': 'display', 'lat': 42.37551146700839...	188	02129	US	Charlestown	MA	United States	[51 Austin Street, Charlestown, MA 02129, Unit...	51f90946498e7e4b05685dff
3	Monument Restaurant & Tavern	Gastropub	251 Main St	School	42.376865	-71.066053	[{'label': 'display', 'lat': 42.37686494887256...	314	02129	US	Charlestown	MA	United States	[251 Main St (School), Charlestown, MA 02129, ...	586d78f9de0cbc0b7f87e131
4	CorePower Yoga	Yoga Studio	19 Austin St Ste E101	NaN	42.375876	-71.065424	[{'label': 'display', 'lat': 42.37587569999999...	192	02129	US	Charlestown	MA	United States	[19 Austin St Ste E101, Charlestown, MA 02129,...	5d60f43992ab450008e007f1
5	Charlestown Liquors	Liquor Store	94-100 Main St	NaN	42.375095	-71.063915	[{'label': 'display', 'lat': 42.37509538273349...	87	02129	US	Charlestown	MA	United States	[94-100 Main St, Charlestown, MA 02129, United...	4b5379f4f964a520e99e27e3
6	Tangierino	Moroccan Restaurant	83 Main St	at Monument Ave.	42.373950	-71.062880	[{'label': 'display', 'lat': 42.37395, 'lng': ...	124	02129	US	Charlestown	MA	United States	[83 Main St (at Monument Ave.), Charlestown, M...	4a0c3163f964a52016751fe3
7	The Battle Of Bunker Hill Museum	History Museum	Monument Square	NaN	42.375772	-71.061159	[{'label': 'display', 'lat': 42.37577204134754...	302	NaN	US	Charlestown	MA	United States	[Monument Square, Charlestown, MA, United States]	4bffefa9daf9c9b6e2f2faef
8	Bunker Hill Monument	Monument / Landmark	31 Monument Sq	NaN	42.376072	-71.060835	[{'label': 'display', 'lat': 42.37607181402598...	342	02129	US	Boston	MA	United States	[31 Monument Sq, Boston, MA 02129, United States]	4a7a5311f964a52029e91fe3
9	Starbucks	Coffee Shop	1 Thompson Sq	NaN	42.375207	-71.064010	[{'label': 'display', 'lat': 42.37520725862152...	97	02129	US	Charlestown	MA	United States	[1 Thompson Sq, Charlestown, MA 02129, United ...	5a7454df6336be6665f8ac62

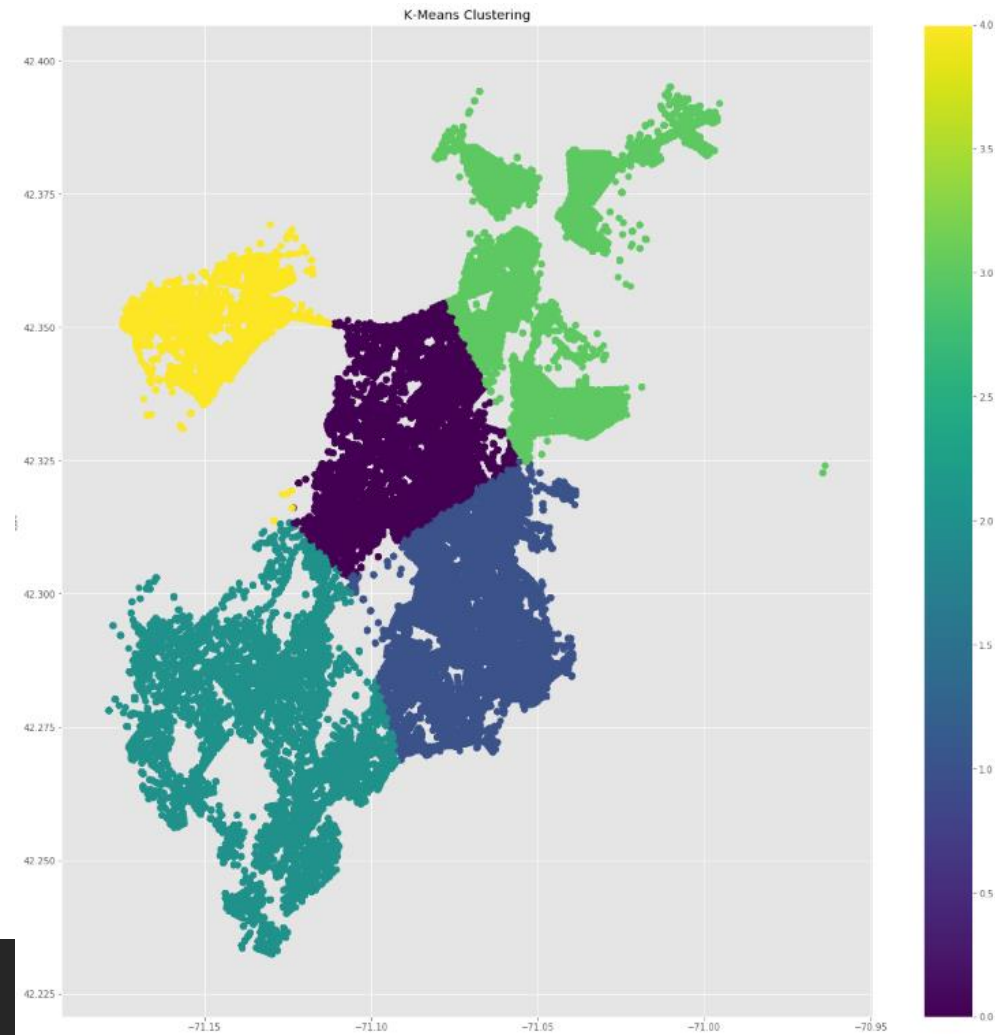# Common Venues in the Safest Area – A15 (Charlestown)

Crime into Two Clusters from 2015 to 2018

Crime into Four Clusters from 2015 to 2018

K-Means Clustering

Crime into Five Clusters from 2015 to 2018

Conclusion

- Key Points:
  - Most crime occurs is at B2, D4, C11, B3, A1
  - Certain venues were closely linked to drug problems and violent crime
  - Motor-Vehicle accident response were the highest number of report registered with the Boston Police.
  - The k-means clustering of the crime data from previous years has indicated that the north part of the police districts have less crime.

- This analysis can help Boston Police act accordingly and try to reduce the crimes frequently occurring in the city of Boston.
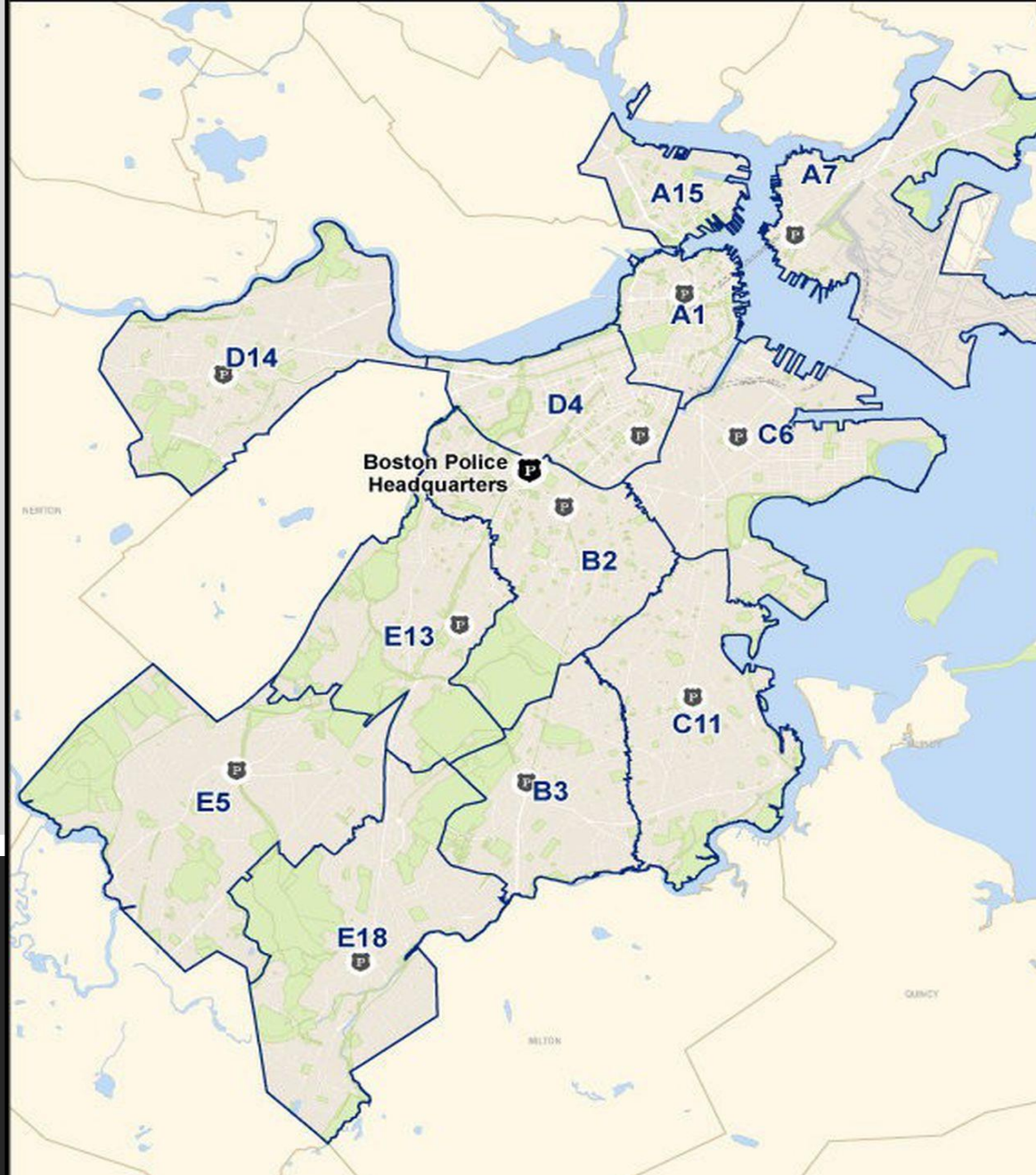
Appendix
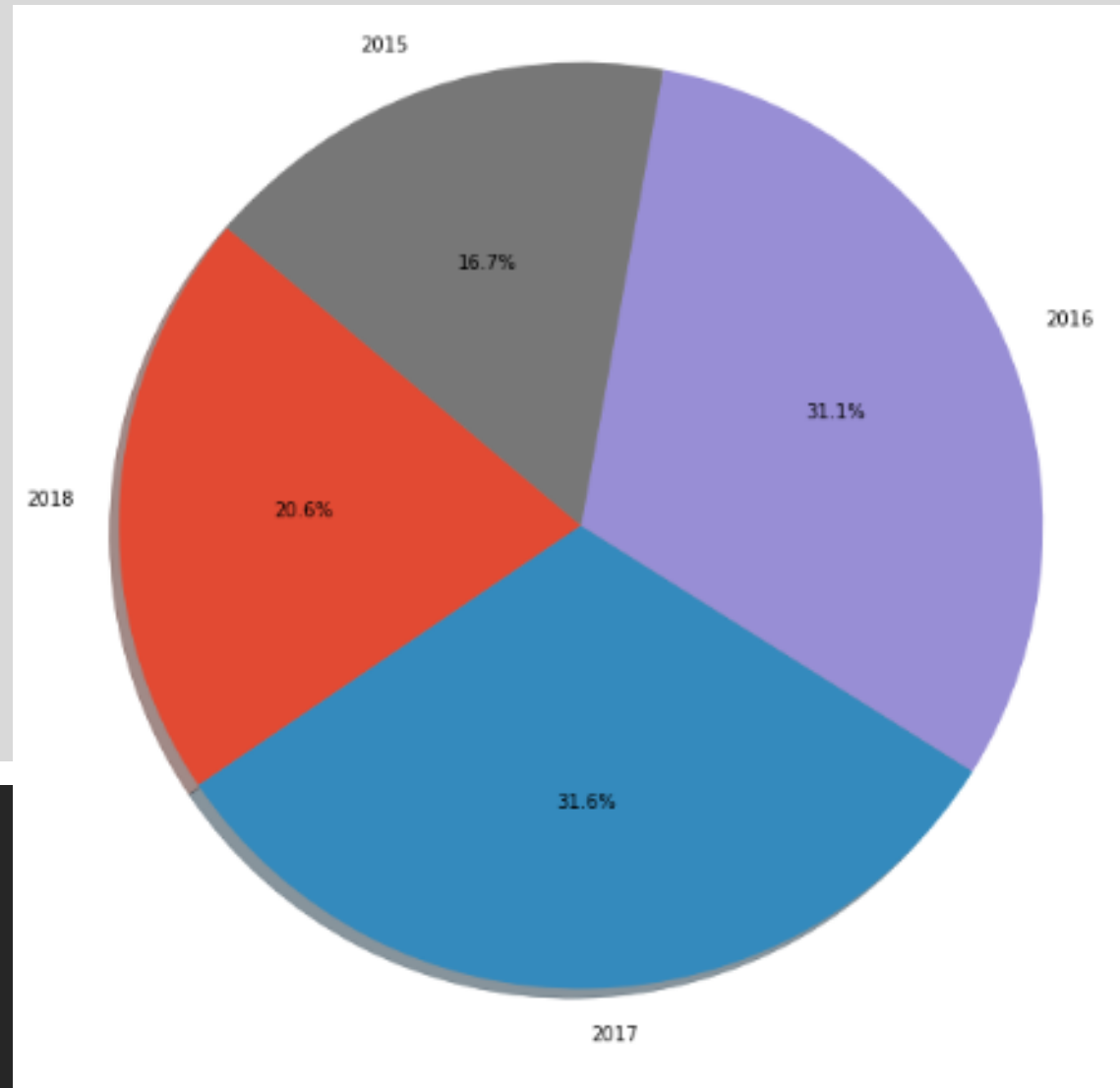
Legend: Current Boston Police Districts Map

This report will also make reference to the Department's various police

districts The districts include A1(Downtown, Beacon Hill, North End, and Chinatown), A7 (East

Boston),A15 (Charlestown), B2 (Roxbury and Mission Hill), B3 (Mattapan and portions of Dorchester), C6 (South Boston), C11 (Dorchester), D4 (Back Bay, South End, and Fenway), D14 (Allston and Brighton), E5 (West Roxbury and Roslindale), E13 (Jamaica Plain), and E18 (Hyde Park).

Current Boston Police Districts Map

The Percentage of Crime from 2015 to 2018

Thank you

- DARKO KNEZEVIC