ANALYZING CRIMES IN US BALTIMORE CITY

Background

- Baltimore is the most populous city in the state of Maryland and the 30th most populous city in the United States, with a population of 593,490 in 2019
- the city is divided into nine geographical regions: North, Northeast, East, Southeast, South, Southwest, West, Northwest, and Central, with each district patrolled by a respective Baltimore Police Department.
- Baltimore is notorious high crime rate.
- Violent crime spiked in 2015 after the death of Freddie Gray on April 19, 2015, which touched off riots and an increase in murders.
- recorded a total of 2027 violent crimes and 4928 property crimes in 2017.
- studying crime dynamics in Baltimore and deriving useful tips will help in reducing crime infestation in the city and suggesting better ways of coping with it

Problem

The objectives of the project are to:

- determine the safest and the most dangerous districts and neighborhoods in Baltimore,
- establish yearly crime trends from 2014 to 2020
- determine crime incidents per each district
- identify the commonest crimes in Baltimore
- locate safest venues in Baltimore

Data

1. BPD Part 1 Victim Based Crime Data

- at: https://data.baltimorecity.gov/Public-Safety/BPD-Part-1-Victim-Based-Crime-Data/wsfq-mvij

2. BPD Districts

at: https://data.baltimorecity.gov/dataset/BPD-Districts/t8n4-my4m

3. Foursquare API

at: https://foursquare.com/explore

Data Science Tools Used

- Python 3.7.6 as the main programming language
- Jupyter notebook 6.0.3 as the execution environment
- Kmeans algorithm for clustering
- Folium for geospatial mapping
- Github to share the project

Explanatory Data Analysis

Acquiring, Exploring and Describing the Data

- The data has 307395 rows and 16 columns

	Longitude	Latitude	Location 1	Total Incidents
count	303902.000000	303902.000000	0.0	307395.0
mean	-76.617477	39.306742	NaN	1.0
std	0.043738	0.029643	NaN	0.0
min	-81.529189	37.577262	NaN	1.0
25%	-76.649136	39.288176	NaN	1.0
50%	-76.614040	39.302721	NaN	1.0
75%	-76.586992	39.326927	NaN	1.0
max	-76.383228	39.661332	NaN	1.0

```
1 baltimore.shape
```

```
(307395, 16)
```

Data wrangling

- changing null weapon to no weapon
- Converting CrimeDate column to date time format
- Extracting days, weeks, weekdays, months and years from CrimeDate

Data wrangling continued..

```
1 baltimore.head()
```

	CrimeDate	CrimeTime	CrimeCode	Location	Description	Inside/Outside	Weapon	Post	District	Neighborhood	Longitude	Latitude	Locatio
0	2020-07-04	10:26:00	4E	1100 N MOUNT ST	COMMON ASSAULT	I	NO WEAPON	724	WESTERN	SANDTOWN-WINCHESTER	-76.644567	39.301295	Nal
1	2020-07-04	06:00:00	6E	2500 KEYWORTH AVE	LARCENY	O	NO WEAPON	533	NORTHERN	GREENSPRING	-76.659229	39.334650	Nal
2	2020-07-04	02:30:00	6D	3300 BRENDAN AVE	LARCENY FROM AUTO	O	NO WEAPON	432	NORTHEAST	BELAIR-EDISON	-76.567938	39.320964	Nal
3	2020-07-04	00:00:00	6D	RD & FRANKFORD AV	LARCENY FROM AUTO	O	NO WEAPON	443	NORTHEAST	NaN	-76.540424	39.320919	Nal
4	2020-07-04	05:40:00	7A	3600 CLIPPER MILL RD	AUTO THEFT	O	NO WEAPON	531	NORTHERN	HAMPDEN	-76.641382	39.331230	Nal

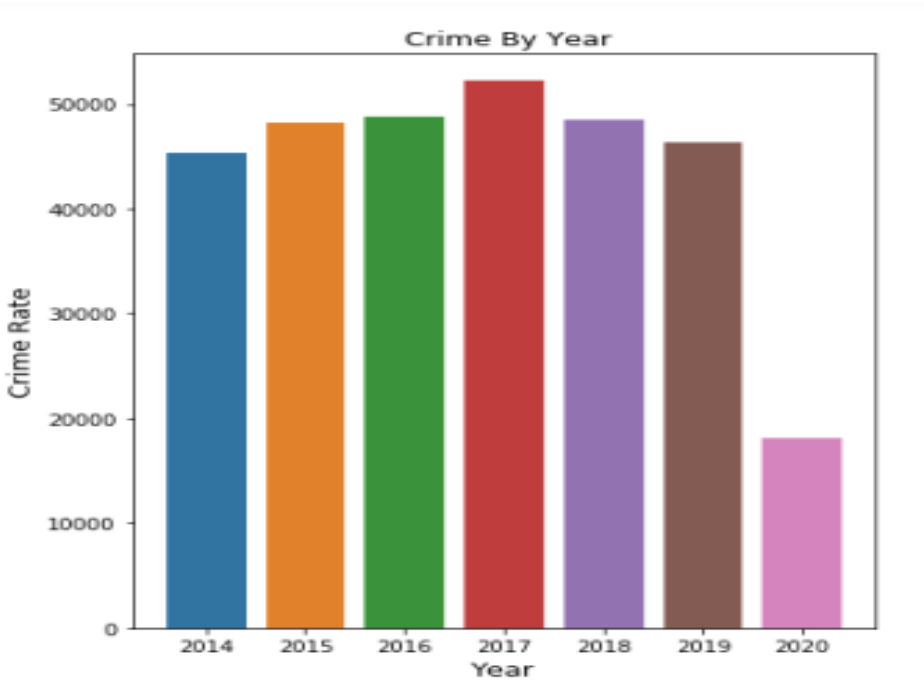
Inside/Outside	Weapon	Post	District	Neighborhood	Longitude	Latitude	Location 1	Premise	vri_name1	Total Incidents	Day	Month	Year	Weekday
I	NO WEAPON	724	WESTERN	SANDTOWN-WINCHESTER	-76.644567	39.301295	NaN	OTHER - INSIDE	NaN	1	4	7	2020	6
O	NO WEAPON	533	NORTHERN	GREENSPRING	-76.659229	39.334650	NaN	STREET	NaN	1	4	7	2020	6
O	NO WEAPON	432	NORTHEAST	BELAIR-EDISON	-76.567938	39.320964	NaN	STREET	Northeastern	1	4	7	2020	6
O	NO WEAPON	443	NORTHEAST	NaN	-76.540424	39.320919	NaN	STREET	NaN	1	4	7	2020	6
O	NO WEAPON	531	NORTHERN	HAMPDEN	-76.641382	39.331230	NaN	DRIVEWAY	NaN	1	4	7	2020	6

Data Analysis

yearly crime incidents from 2014-2020

- 2017 recorded highest crimes while 2020 has fewest crimes because the year is just beginning

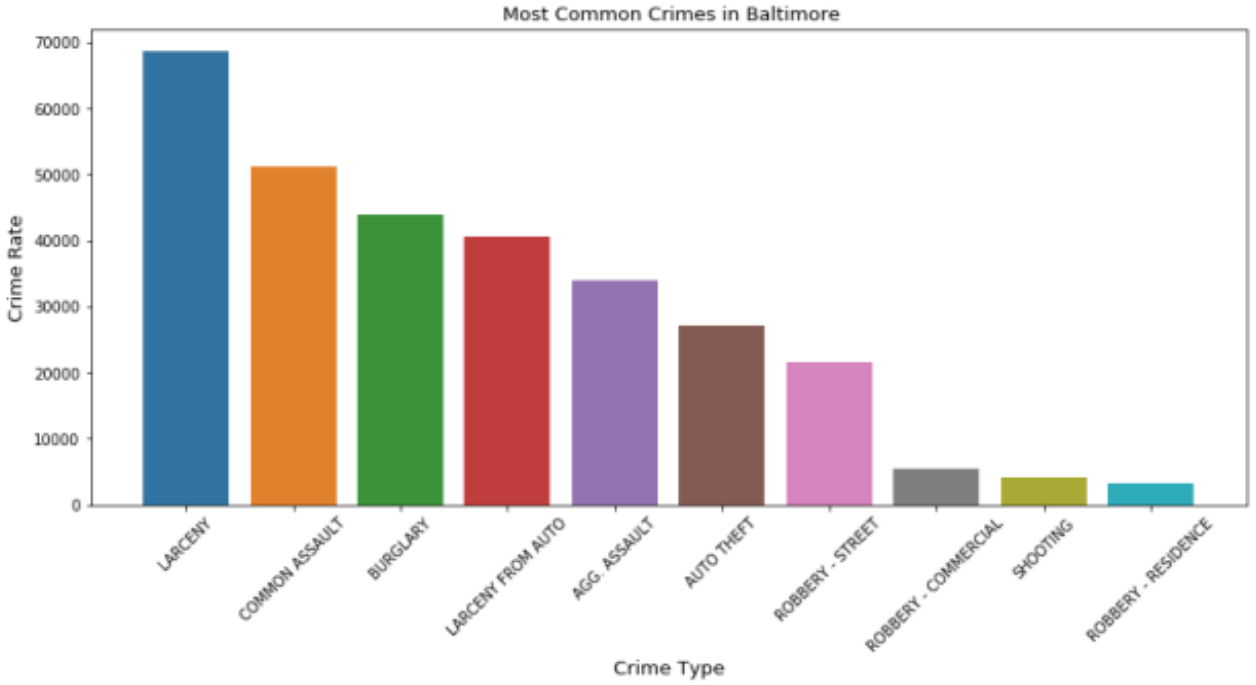
	year	crimes
0	2017	52174
1	2016	48782
2	2018	48468
3	2015	48192
4	2019	46312
5	2014	45309
6	2020	18071



10 most common crimes in Baltimore

- Larceny (a type of theft) is the commonest while residential robbery is the least on the first ten crimes in Baltimore list

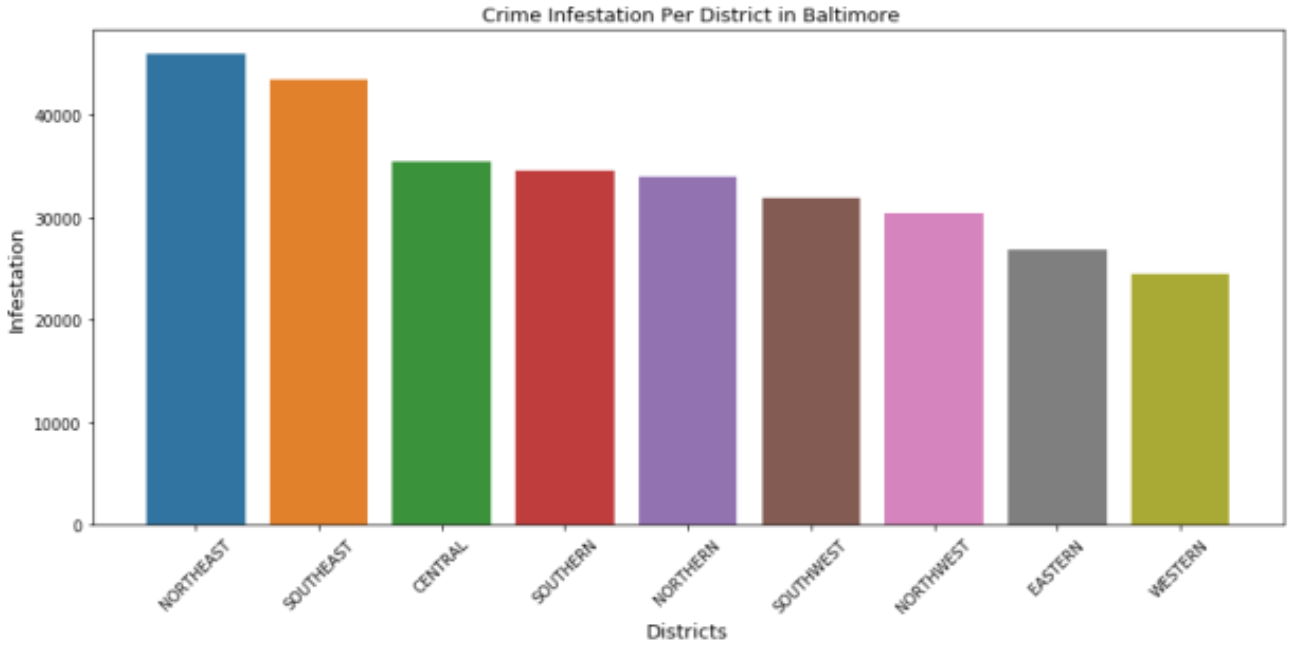
	crime	rate
0	LARCENY	68577
1	COMMON ASSAULT	51096
2	BURGLARY	43921
3	LARCENY FROM AUTO	40468
4	AGG. ASSAULT	33830
5	AUTO THEFT	26994
6	ROBBERY - STREET	21650
7	ROBBERY - COMMERCIAL	5427
8	SHOOTING	4133
9	ROBBERY - RESIDENCE	3212



Crimes per Region

- North East is the worst hit district while Western is the safest

	region	rate
0	NORTHEAST	45940
1	SOUTHEAST	43380
2	CENTRAL	35462
3	SOUTHERN	34526
4	NORTHERN	33949
5	SOUTHWEST	31843
6	NORTHWEST	30348
7	EASTERN	26783
8	WESTERN	24545



Kmeans Clustering

- Clustering districts into three groups of highest, lowest and average crime hit.
- Kcluster =3
- Data used for clustering is crime rates
- The Output of kmeans was an array of [1,1,0,0,0,0,0,2,2]
- Kmeans output: array ([1, 1, 0, 0, 0, 0, 0, 2, 2])

	rate
0	45940
1	43380
2	35462
3	34526
4	33949
5	31843
6	30348
7	26783
8	24545

RESULTS

Clustered Districts:

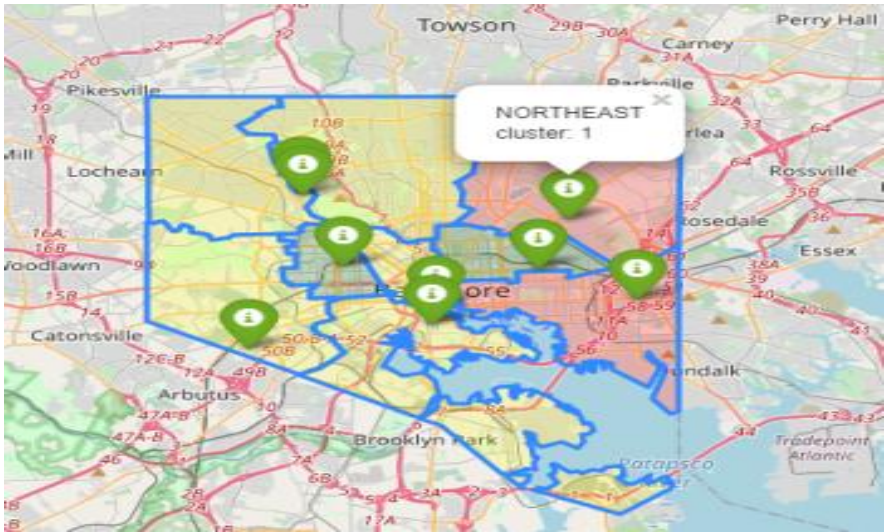
- The districts and their respective clusters
- Cluster 1 for the highest crime regions, cluster 2 for regions with lowest crime numbers, and cluster 0 for average crime rates.

	region	rate	cluster	Longitude	Latitude
0	NORTHEAST	45940	1	-76.567938	39.320964
1	SOUTHEAST	43380	1	-76.544564	39.287601
2	CENTRAL	35462	0	-76.613466	39.285077
3	SOUTHERN	34526	0	-76.614368	39.276819
4	NORTHERN	33949	0	-76.659229	39.334650
5	SOUTHWEST	31843	0	-76.677057	39.266628
6	NORTHWEST	30348	0	-76.658445	39.331355
7	EASTERN	26783	2	-76.578083	39.300148
8	WESTERN	24545	2	-76.644567	39.301295

Results…

Mapping Clustered Regions of Baltimore:

- The clustered districts mapped in three colors:
  - Red: dangerous districts, cluster 1
  - Green: safest districts, cluster 2
  - Yellow: average districts, cluster 0



Results…

Safest Neighborhoods in Baltimore

- The safest neighborhoods were derived from safest districts.
- There are 48977, relatively, safest neighborhoods in Baltimore as of date of this analysis.

	Neighborhood	Longitude	Latitude
0	SANDTOWN-WINCHESTER	-76.644567	39.301295
6	HARLEM PARK	-76.639986	39.295319
12	MADISON EASTEND	-76.578083	39.300148
28	MOSHER	-76.662150	39.298075
37	SOUTH CLIFTON PARK	-76.589462	39.312452

```
1 safe_Neighborhoods.shape
```

```
(48977, 3)
```

Results ….

Using Foursquare to Get Venues of Harlem Park

- Harlem Park is one of the safest neighborhoods as per the analysis.
- only five venues in Harlem Park neighborhood.

	venue	category	lat	lng
0	Shareef's Grill	American Restaurant	39.294184	-76.638196
1	Lafayette Square Park	Park	39.298028	-76.635910
2	Harlem Park	Park	39.296098	-76.641376
3	Harlem Park Pl	Playground	39.296684	-76.640016
4	The Hide Out	Bar	39.295032	-76.644929

```
1 nearby_venues.shape
```

```
(5, 4)
```

	Neighborhood	Longitude	Latitude	venue	category	lat	lng
0	HARLEM PARK	-76.639986	39.295319	Shareef's Grill	American Restaurant	39.294184	-76.638196
1	HARLEM PARK	-76.633312	39.296943	Lafayette Square Park	Park	39.298028	-76.635910
2	HARLEM PARK	-76.632678	39.295967	Harlem Park	Park	39.296098	-76.641376
3	HARLEM PARK	-76.646465	39.297514	Harlem Park Pl	Playground	39.296684	-76.640016
4	HARLEM PARK	-76.645779	39.296493	The Hide Out	Bar	39.295032	-76.644929

DISCUSSION

- Baltimore city has a crime problem.
- used the crime data provided by the Baltimore police department to establish the most common crimes in Baltimore, yearly crime progresses since 2014, crime figures per district and to cluster the regions/districts into 3 blocks based on crime infestation using Kmeans.
- also identified safest neighborhoods and safest venues in one of these safe neighborhoods.

summary of the findings:

- From 2014-2020, 2017 recorded highest crime incidents while 2020 recorded fewest crime incidents
- The most common crimes in Baltimore are larceny, assaults, burglary, theft, robbery and shootings. This goes against the popular belief about Baltimore being famed for police shootings and homicides as reported more often by the media.
- Across the districts, Northeast is the worst hit in terms of crimes while Western had least crime incidents compared to other districts.
- From the clustering, Northeast and Southeast are the most dangerous districts in Baltimore while Eastern and Western are the safest. The rest remained at average.
- Eastern and Western have a total of 48977 neighborhoods considered safest neighborhoods in Baltimore.
- Harlem Park has five venues.

CONCLUSION

- I hope my findings will be seen useful by investors, businesspeople, tourists, police and others living in or coming to Baltimore.
- If the findings help in tackling Baltimore crime problem in one way or another, then the goal of the study has been achieved and machine learning analysis justified.
- However, more analytical works need to be done on Baltimore crime issue to get to the root cause of the problem.