Chicago City Neighborhood Crime Analysis

Shafiul Azam

Introduction

Motivation of the Study

- Chicago has violent crime about 49% higher than the average rate of crime in Illinois and 164% higher than the rest of the nation (https://www.areavibes.com/chicago-il/crime/).
- Therefore, it is often important for any potential home buyers or renters do some study and analyze the crime data before making decisions of a safe neighborhood.

DATA DESCRIPTION

1. Neighborhood Information was available from the following wiki page: (https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Chicago)

2. Crime information was available from Chicago Data Portal

(https://data.cityofchicago.org/Public-Safety/Crimes-Map/dfnk-7re6).

Elements in Crime Reports

Case #
Date of Occurrence
Approximate Block Address.
Primary Description of the crime (i.e. Burglary, Battery, Theft etc.)
Secondary Description of the Crime
Location Description
Arrest made (Y/N)?
Domestic Violence (Y/N)?
GPS Point (Latitude and Longitude)

Methodology

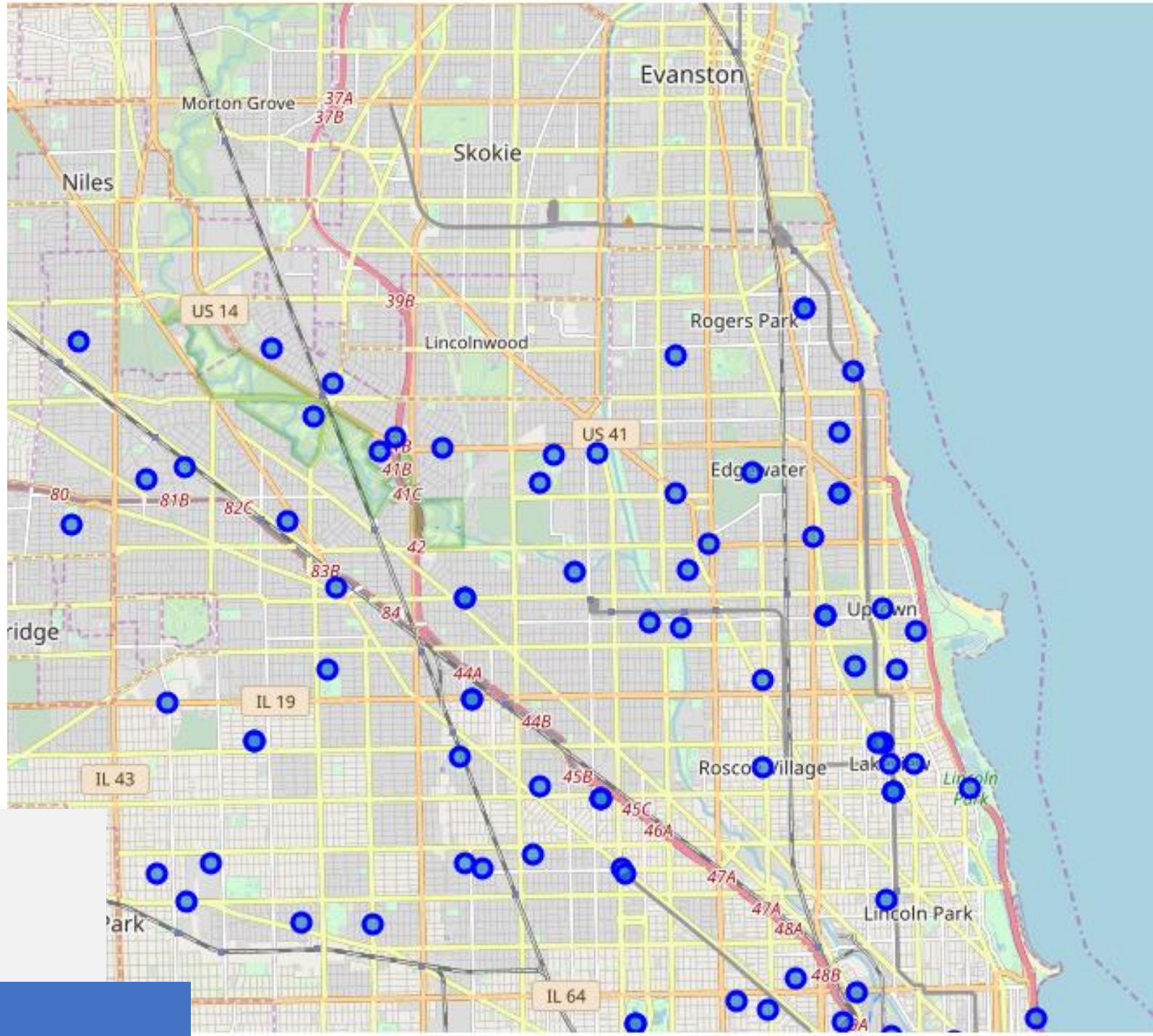
Mapping the Neighbor

Crime Data Import and Data Wrangling

Exploratory Analysis of Crime Data

Clustering of Neighborhoods using K-means clustering

Chicago Neighborhoods



Data Wrangling

1

Step 1 - Import Crime Data from City Portal

2

Step 2- Preparation of Data for Crime Distribution

3

Step 3 – Data Preparation for Temporal Analysis

4

Step 4 – Calculate Neighborhood Coordinates using geopy package

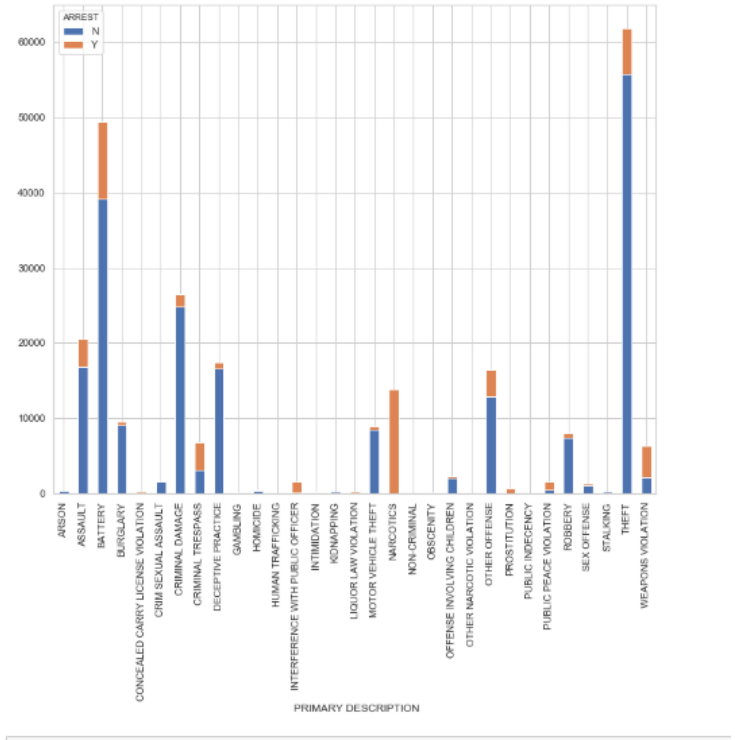
5

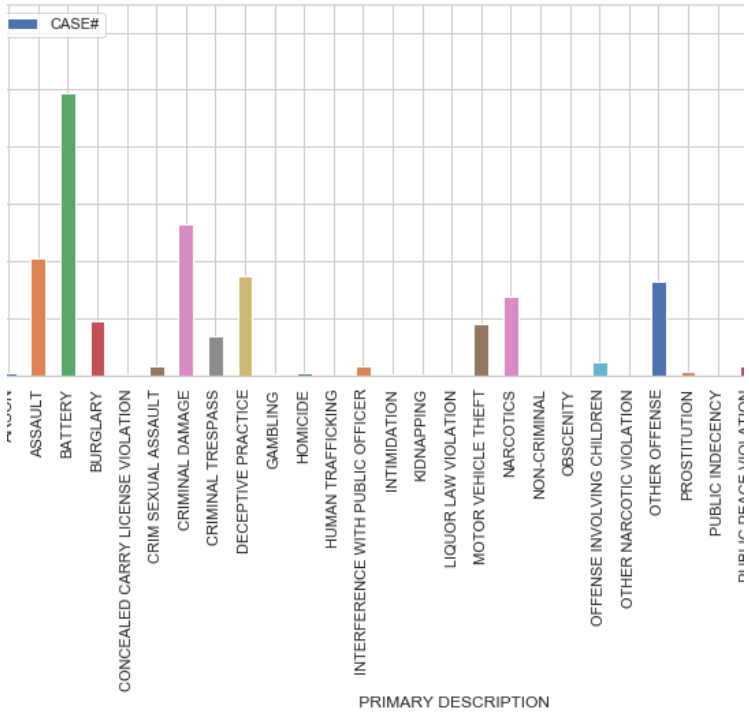
Step 5 – Develop functions to link crime to neighborhood by nearest location

6

Step 6 – One hot coding to create dummy variable for crime types

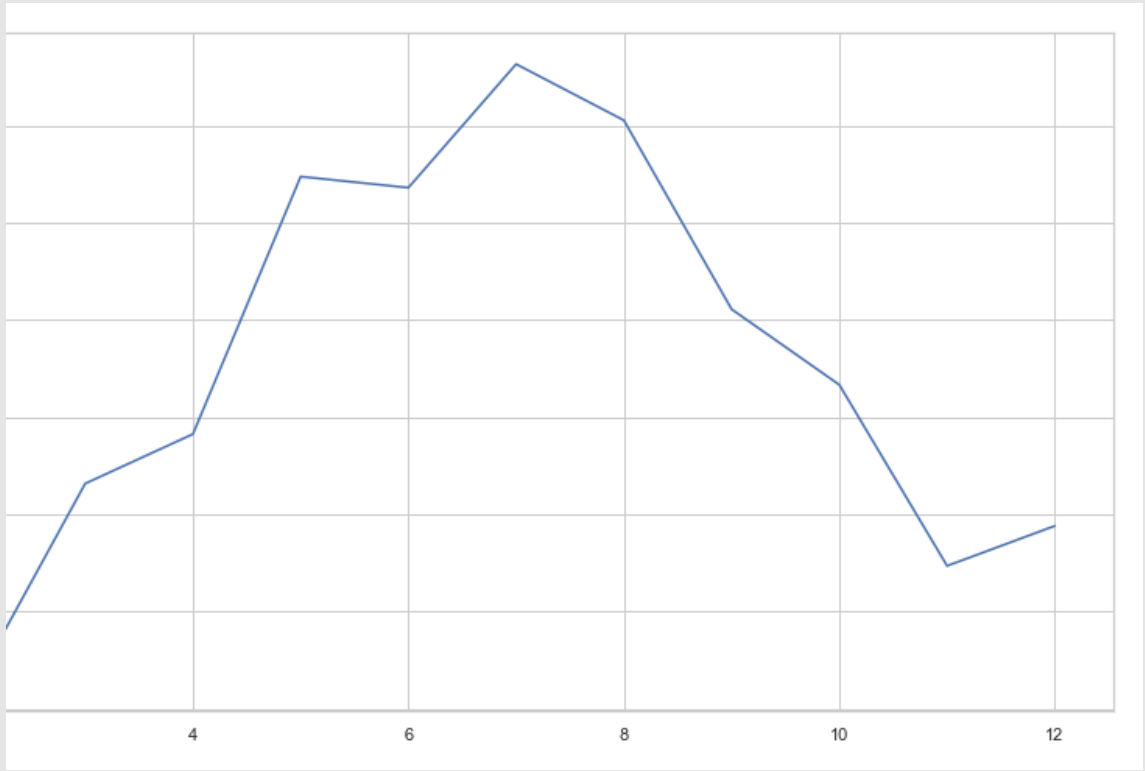
Exploratory Analysis: Distribution of Crime Types in Chicago





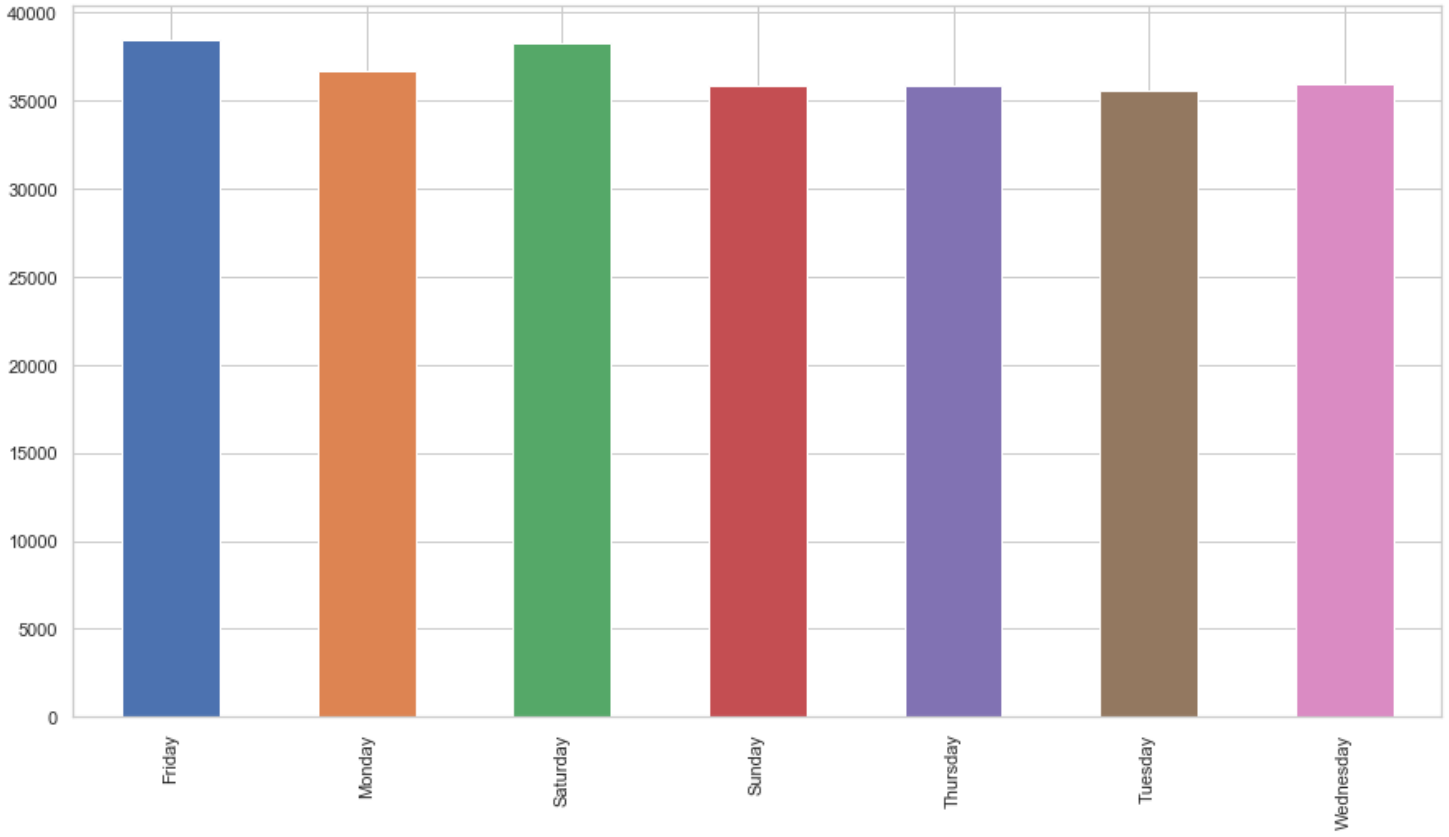
- It is evident that almost all Narcotics related crime ended up having arrest.
- However, for battery or theft the percentage of arrests being made are pretty low.

Temporal Analysis



- The months of July and August have typically more crimes.

Crime by Days of Week

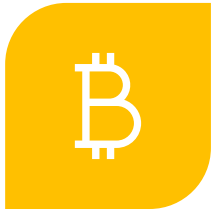


The distribution of crime by day of weeks do not show any noticeable differences

K-neighborhood clustering

THE BASIS OF THE K-NEIGHBORHOOD ANALYSIS WAS TO CRIME TYPES SO THAT NEIGHBORHOOD WITH SIMILAR CRIME TRENDS CAN BE CLUSTERED TOGETHER.

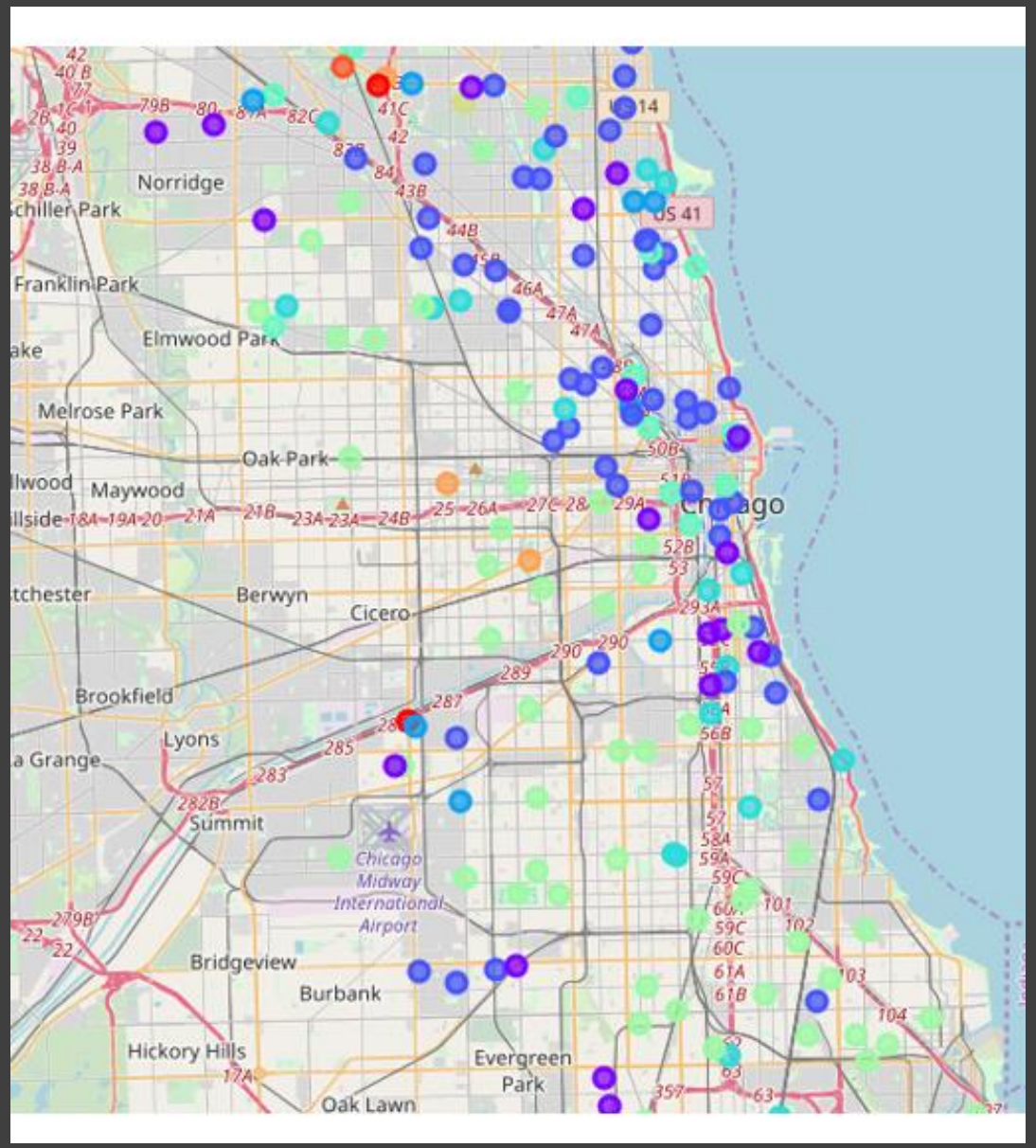
BEFORE APPLYING THE K-NEIGHBORS CLASSIFIER, THE CRIME DATA WAS USED TO DISPLAY THE DISTRIBUTION OF THE CRIME ACROSS ANY NEIGHBORHOOD.



FOR EXAMPLE, FOR ANDERSONVILLE NEIGHBORHOOD THE DISTRIBUTION OF THE CRIME IS SHOWN HERE:

----Andersonville----

	Neighborhood	freq
0	THEFT	0.33
1	OTHER OFFENSE	0.13
2	CRIMINAL DAMAGE	0.13
3	BATTERY	0.10
4	BURGLARY	0.07
5	CRIMINAL TRESPASS	0.07
6	MOTOR VEHICLE THEFT	0.07
7	ASSAULT	0.07
8	DECEPTIVE PRACTICE	0.03
9	ARSON	0.00



K-neighborhood clustering technique.

- Finally, K value of 10
- The map displays the identified clusters

Results and DIscussion

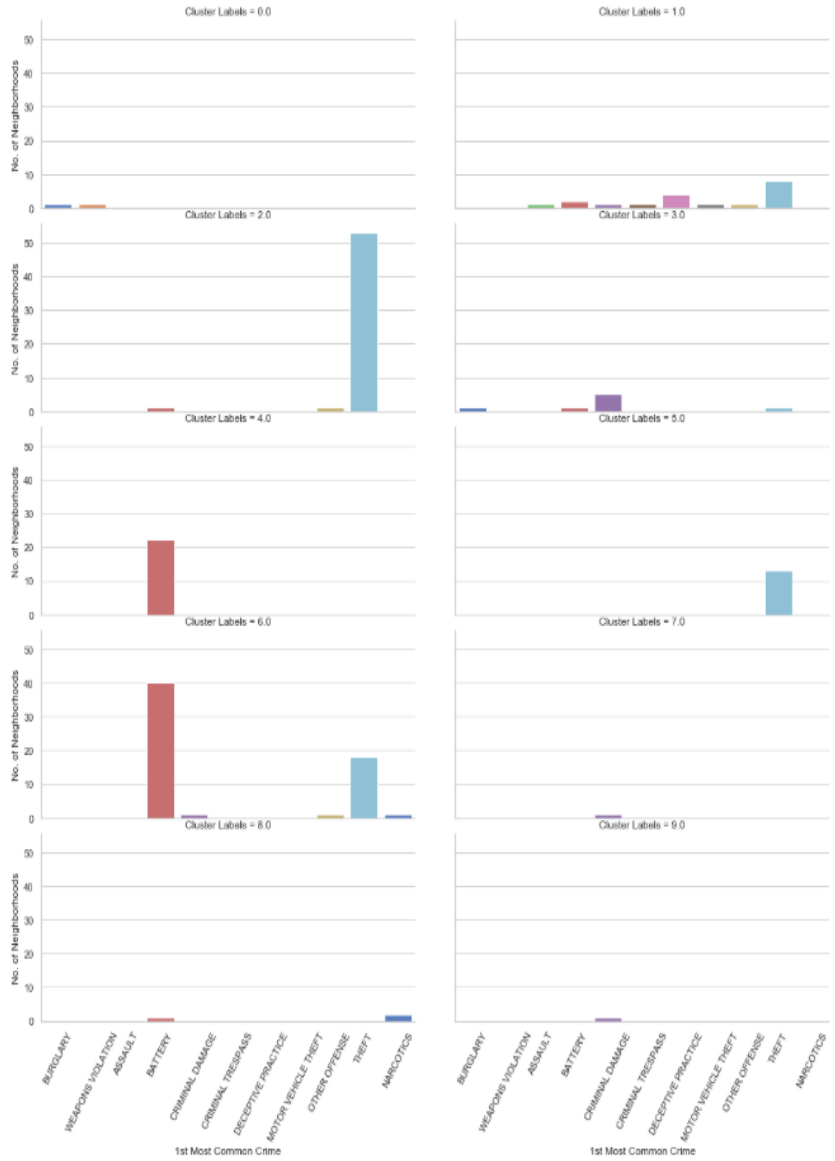
	Community Areas	1st Most Common Crime	2nd Most Common Crime	3rd Most Common Crime	4th Most Common Crime	5th Most Common Crime	6th Most Common Crime	7th Most Common Crime	8t
111	Garfield Ridge	WEAPONS VIOLATION	BURGLARY	CRIMINAL DAMAGE	OTHER OFFENSE	KIDNAPPING	ASSAULT	BATTERY	
204	Forest Glen	BURGLARY	OTHER OFFENSE	WEAPONS VIOLATION	KIDNAPPING	ASSAULT	BATTERY	CONCEALED CARRY LICENSE VIOLATION	

Individual Clusters are obtained with the Top 10 Most Common Crimes. For example, the cluster 1 and 3 has the following crime distributions.

	Community Areas	1st Most Common Crime	2nd Most Common Crime	3rd Most Common Crime	4th Most Common Crime	5th Most Common Crime	6th Most Common Crime	7th Most Common Crime	8th Most Common Crime	9th Most Common Crime	Co
2	Edgewater	THEFT	CRIMINAL DAMAGE	OTHER OFFENSE	BATTERY	ASSAULT	BURGLARY	CRIMINAL TRESPASS	MOTOR VEHICLE THEFT	DECEPTIVE PRACTICE	
3	Archer Heights	THEFT	CRIMINAL DAMAGE	ASSAULT	BATTERY	OTHER OFFENSE	BURGLARY	CRIMINAL TRESPASS	DECEPTIVE PRACTICE	MOTOR VEHICLE THEFT	
5	Ashburn	THEFT	OTHER OFFENSE	ASSAULT	BATTERY	MOTOR VEHICLE THEFT	INTERFERENCE WITH PUBLIC OFFICER	CRIMINAL DAMAGE	OFFENSE INVOLVING CHILDREN	WEAPONS VIOLATION	
9	Avondale	THEFT	BATTERY	DECEPTIVE PRACTICE	CRIMINAL DAMAGE	ASSAULT	OTHER OFFENSE	ROBBERY	MOTOR VEHICLE THEFT	CRIMINAL TRESPASS	
12	Belmont Cragin	THEFT	DECEPTIVE PRACTICE	CRIMINAL DAMAGE	BATTERY	BURGLARY	MOTOR VEHICLE THEFT	OTHER OFFENSE	ROBBERY	NARCOTICS	

Distribution of clusters by 1st Most Common Crime

- There are diverse distribution of crime rates in the neighborhood.
- Cluster#2 has around 53 neighborhoods that has Theft pretty common.
- Cluster#6 has battery very common
- Cluster#3 has significant Criminal damages.



Conclusion

- The study shows an application of web scraping, geo-coding and unsupervised machine learning technique (i.e. K-nearest neighborhood analysis) to analyze and cluster residential neighborhood in terms of types and extensity of crimes.
- The results can help a home buyer or renter choosing a safer neighborhood or cluster.
- It can also help law enforcement agencies consider strategic planning to mitigate the crimes.