The Battle of Neighborhoods

Popular Venue and neighborhood crime in Uptown Chicago

This is the final report on capstone project. For reference Original requirements of report submission mention below which is:

- A full report consisting of all of the following components (**15 marks**):

- **Introduction** where you discuss the business problem and who would be interested in this project.
- **Data** where you describe the data that will be used to solve the problem and the source of the data.
- **Methodolog**y section which represents the main component of the report where you discuss and describe any exploratory data analysis that you did, any inferential statistical testing that you performed, if any, and what machine learnings were used and why.
- **Result**s section where you discuss the results.
- **Discussion** section where you discuss any observations you noted and any recommendations you can make based on the results.
- **Conclusion** section where you conclude the report.

Introduction

1.1Background

City of Chicago

City of Chicago is one of the most famous city in US. It's mostly known for its iconic landmarks, wide range diversity and so on. It is the most populous city in state of IL and 3rd most populous city in US according to last year (WIKI, n.d.).It is an international hub for financial activity, education, technology, telecommunication, culture and commerce. Chicago city is divided into 77 community by the researcher of university of Chicago in the year 1920 and these communities are divided into more than 200 neighborhoods .City wise its population is evenly distributed among Hispanic, Black, White, and Asian. At micro label its neighborhood is quite segregated. Diversity of population in a community or neighborhood reflects on the types of available ethnic restaurant. Chicago city it also known for its high violent crime rate among the nation. So there is also some fear always come across through the travelers minds who wants to explore any not so popular neighborhood.

Chicago Uptown Community

Chicago uptown community is historically famous in north side of the city. It boast the city's one of the diverse community because of its welcoming nature towards the immigrants from all corn -er of the worlds. Uptown's boundaries are Foster Avenue on the north; Lake Michigan on the east; Montrose (Ravenswood to Clark), and Irving Park (Clark Street to Lake Michigan) on the south; Ravenswood (Foster to Montrose), and Clark (Montrose to Irving Park) on the west. To the north is Edgewater, to the west is Lincoln Square, and to the south is Lake View. It is divided into 5 neighborhood. They are

- Buena Park,
- Clarendon Park
- Margate Park
- New Chinatown/Argyle Park
- Sheridan Park
- Uptown

1.2 Business problem

Few public place like Chicago Lakeshore Hospital, Methodist Hospital of Chicago, Thorek Memorial, Hospital, Northwestern University, City college of Chicago are parts of or nearby this community. That means lots of people visit these neighborhoods from somewhere else daily basis. Suppose some grad student came to northwestern university campus to attain a whole day conference or job fare from other part of IL or Chicago. No matter how busy his schedule is he needs to give some fuel to his body i.e food. Since uptown is also famous for its ethnic restaurant, so he is curious about which ethnic restaurants are dominant around this community or are the whole uptown community has same characteristic or they have some dissimilarity. Again few Chicago area are well known for its high crime rate. So this project will do some exploratory crime data analysis.

1.3 who would be interested in this project?

If you are planning to have trip to Northside Chicago soon, than this project is plan and executed for you. Before reached to your destination if you get chance to read this report than BAM!!! You already know where to go have your lunch /dinner or simply a cup of hot coffee without any fear of discovering you self in a violent situation.

Data

2.1 Data Description

In this section you will get the information about type and source of data and how the data help to solve the problem.

- I use open source Chicago Crime data to provide the user with additional crime data. The data can be found through the following link: https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-present/ijzp-q8t2

- Chicago uptown neighborhood and community data is scraped from the wiki page. You can found the wiki page here: https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Chicago

- Latitude and longitude of the all the neighborhood was searched in here: https://www.latlong.net/

- I use FourSquare API to get supplemental geographical data about the neighborhood venue

2.2 Data Cleaning

Chicago Crime Data

This dataset reflects reported incidents of crime (with the exception of murders where data exists for each victim) that occurred in the City of Chicago from 2001 to present, minus the most recent seven days. Data is extracted from the Chicago Police Department's CLEAR (Citizen Law Enforcement Analysis and Reporting) system. Due to limitation of computational capabilities we downloads only 2019 data.

Next page you will get crime data column information.

Column Name	Type	Description
CASE#	Plain Text	The Chicago Police Department RD Number
DATE OF OCCURRENCE	Date & Time	Date when the incident occurred.
BLOCK	Plain Text	The partially redacted address where the incident occurred, placing it on the same block as the actual address.
IUCR	Plain Text	The Illinois Unifrom Crime Reporting code. This is directly linked to the Primary Type and Description.
PRIMARY DESCRIPTION	Plain Text	The primary description of the IUCR code.
SECONDARY DESCRIPTION	Plain Text	The secondary description of the IUCR code, a subcategory of the primary description.
LOCATION DESCRIPTION	Plain Text	Description of the location where the incident occurred.
ARREST	Plain Text	Indicates whether an arrest was made.
DOMESTIC	Plain Text	Indicates whether the incident was domestic-related as defined by The Illinois Domestic Violence Act.
BEAT	Plain Text	Indicates the beat where the incident occurred. A beat is the smallest police geographic area – each beat has a dedicated police beat car. Three to five beats make up a police sector, and three sectors make up a police district. The Chicago Police Department has 22 police districts. See the beats at https://data.cityofchicago.org/d/aerh-rz74.
WARD	Number	The ward (City Council district) where the incident occurred. See the wards at https://data.cityofchicago.org/d/sp34-6z76.
FBI CD	Plain Text	Indicates the crime classification as outlined in the FBI's National Incident-Base

		.
X COORDINATE	Plain Text	The x coordinate of the location where the incident occurred in State Plane Illinois East NAD 1983 projection. This location is shifted from the actual location for partial redaction but falls on the same block.
Y COORDINATE	Plain Text	The y coordinate of the location where the incident occurred in State Plane Illinois East NAD 1983 projection. This location is shifted from the actual location for partial redaction but falls on the same block.
LATITUDE	Number	The latitude of the location where the incident occurred. This location is shifted from the actual location for partial redaction but falls on the same block.
LONGITUDE	Number	The longitude of the location where the incident occurred. This location is shifted from the actual location for partial redaction but falls on the same block.
LOCATION	Location	The location where the incident occurred in a format that allows for creation of maps and other geographic operations on this data portal. This location is shifted from the actual location for partial redaction but falls on the same block.

	CASE#	DATE OF OCCURRENCE	BLOCK	IUCR	PRIMARY DESCRIPTION	SECONDARY DESCRIPTION	LOCATION DESCRIPTION	ARREST	DOMESTIC	BEAT	WARD	FBI CD	X COORDINATE	COORD
0	JC497784	11/03/2019 11:40:00 AM	032XX N CLARK ST	0860	THEFT	RETAIL THEFT	DEPARTMENT STORE	N	N	1924	44.0	06	NaN	
1	JB556584	11/29/2018 10:04:00 AM	002XX W 23RD ST	0810	THEFT	OVER $500	ALLEY	N	N	914	25.0	06	NaN	
2	JC497415	11/03/2019 04:30:00 AM	107XX S PEORIA ST	1320	CRIMINAL DAMAGE	TO VEHICLE	RESIDENTIAL YARD (FRONT/BACK)	N	N	2233	34.0	14	NaN	
3	JB559847	12/19/2018 01:14:00 PM	042XX W WILCOX ST	2014	NARCOTICS	MANU/DELIVER: HEROIN (WHITE)	ALLEY	Y	N	1115	28.0	18	NaN	
4	JB527374	11/23/2018 02:15:00 PM	0000X S STATE ST	0810	THEFT	OVER $500	STREET	N	N	112	42.0	06	NaN	

Now only the attributes listed below are required for our analysis:

- Date of Occurrence
- Primary Description
- Ward
- Latitude
- Longitude

Uptown is divided into multiple wards, which are the districts from which aldermen in the Chicago City Council are drawn. Most of the community area lies in the 46th and 48th wards, with small portions of the neighborhood's west side located in the 47th and 40th wards. So for this project data for ward 40, 46, 47, 48 are used.

Then the data is processed below:

- Cleanup column name
- Change the date of occurrence field to a date / time object
- Add new columns for:

    1. Hour
    2. Day
    3. Month
    4. Month_NO
    5. Year

- Get data frame for ward 40,46,47,48 and then concat to get the crime data for Uptown.

.

In [87]:
```
crime_uptown.head()
```

Out[87]:

	CASE	DATE	DESCRIPTION	WARD	LATITUDE	LONGITUDE	HOUR	MONTH	DAY	YEAR	MONTH_NO
73	JC421371	2019-09-05 12:24:00	DECEPTIVE PRACTICE	46.0	41.962037	-87.645884	12	September	Thursday	2019	9
100	JC497471	2019-11-03 19:52:00	BATTERY	46.0	41.965599	-87.647776	19	November	Sunday	2019	11
103	JC421354	2019-09-05 12:24:00	THEFT	46.0	41.962037	-87.645884	12	September	Thursday	2019	9
306	JC421213	2019-09-05 14:20:00	ASSAULT	46.0	41.969078	-87.655608	14	September	Thursday	2019	9
313	JC424888	2019-09-05 08:00:00	DECEPTIVE PRACTICE	46.0	41.946459	-87.644655	8	September	Thursday	2019	9

Chicago Neighborhood and community Data

Chicago neighborhood and community data table is read as read_htm . This convert html table to data frame object. Here all the table are in a list. Then we get the list element in which our desired table contained.

Getting data about Chicago Neighborhood

```
import pandas as pd # import pandas for scrapping data of chicago Neiborhood and community name from html file
```

```
 pd.__version__
```

```
'0.25.1'
```

```
table = pd.read_html('https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Chicago', index_col= False)
```

```
d =table[0]
```

```
d.columns =[ 'Neighborhood','Community'] # fix the index name with proper name
```

```
d.head()
```

	Neighborhood	Community
0	Albany Park	Albany Park
1	Altgeld Gardens	Riverdale
2	Andersonville	Edgewater
3	Archer Heights	Archer Heights
4	Armour Square	Armour Square

Chicago neighborhood:

We are getting name of all the community of uptown by this code and replacing name Argyle Park with its new name New Chinatown

df.replace({'argyle park':'New Chinatown'},inplace = True)

df =d[d['Community'] =='Uptown'].reset_index(drop=True)

```
In [130]: df.shape
Out[130]: (6, 2)
```

```
In [131]: df.head(6)
Out[131]:
```

	Neighborhood	Community
0	Buena Park	Uptown
1	Clarendon Park	Uptown
2	Margate Park	Uptown
3	New Chinatown	Uptown
4	Sheridan Park	Uptown
5	Uptown	Uptown

Figure: Neighborhood name of uptown community

Longitude latitude for each neighborhood are added as separate column of data frame.

```
35]: df['Latitude'] =Latitude
     df['Longitude'] = Longitude

     df.head()
```

35]:

	Neighborhood	Community	Latitude	Longitude
0	Buena Park	Uptown	41.957810	-87.652833
1	Clarendon Park	Uptown	41.963275	-87.648842
2	Margate Park	Uptown	41.972465	-87.652863
3	New Chinatown	Uptown	41.973400	-87.658600
4	Sheridan Park	Uptown	41.965500	-87.663000

Data from foursquare

We construct the url with our CLIENT_id, CLIENT_secret and VERSION for all the 6 neighborhood longitude and latitude and get top 100 venue within 500 meter radius. Get the request of constructed url in .JSON file.

By the get_catagory_type function we get all venue in each neighborhood

```
# function that extracts the category of the venue
def get_category_type(row):
    try:
        categories_list = row['categories']
    except:
        categories_list = row['venue.categories']

    if len(categories_list) == 0:
        return None
    else:
        return categories_list[0]['name']
```

```
venues = results['response']['groups'][0]['items']

nearby_venues = json_normalize(venues) # flatten JSON

# filter columns
filtered_columns = ['venue.name', 'venue.categories', 'venue.location.lat', 'venue.location.lng']
nearby_venues =nearby_venues.loc[:, filtered_columns]

# filter the category for each row
nearby_venues['venue.categories'] = nearby_venues.apply(get_category_type, axis=1)

# clean columns
nearby_venues.columns = [col.split(".")[-1] for col in nearby_venues.columns]

nearby_venues.head()
```

[145]:

	name	categories	lat	lng
0	Klein's Bakery & Cafe	Bakery	41.958328	-87.652953
1	Michael's Original Pizzeria & Tavern	Pizza Place	41.956879	-87.651865
2	Bar on Buena	Bar	41.958528	-87.653579
3	North Buena Deli and Wine Shop	Wine Shop	41.958474	-87.653173
4	Siam Noodle and Rice	Asian Restaurant	41.957937	-87.652906

```
#Create a function to repeat the same process to all the neighborhoods in uptown Chicago

def getNearbyVenues(names, latitudes, longitudes, radius=500):

    venues_list=[]
    for name, lat, lng in zip(names, latitudes, longitudes):
        print(name)

        # create the API request URL
        url = 'https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&limit={}'.format(
            CLIENT_ID,
            CLIENT_SECRET,
            VERSION,
            lat,
            lng,
            radius,
            LIMIT)

        # make the GET request
        results = requests.get(url).json()["response"]['groups'][0]['items']

        # return only relevant information for each nearby venue
        venues_list.append([(
            name,
            lat,
            lng,
            v['venue']['name'],
            v['venue']['location']['lat'],
            v['venue']['location']['lng'],
            v['venue']['categories'][0]['name']) for v in results])

    nearby_venues = pd.DataFrame([item for venue_list in venues_list for item in venue_list])
    nearby_venues.columns = ['Neighborhood',
                  'Neighborhood_Latitude',
                  'Neighborhood_Longitude',
                  'Venue',
                  'Venue_Latitude',
                  'Venue_Longitude',
                  'Venue_Category']

    return(nearby_venues)
```

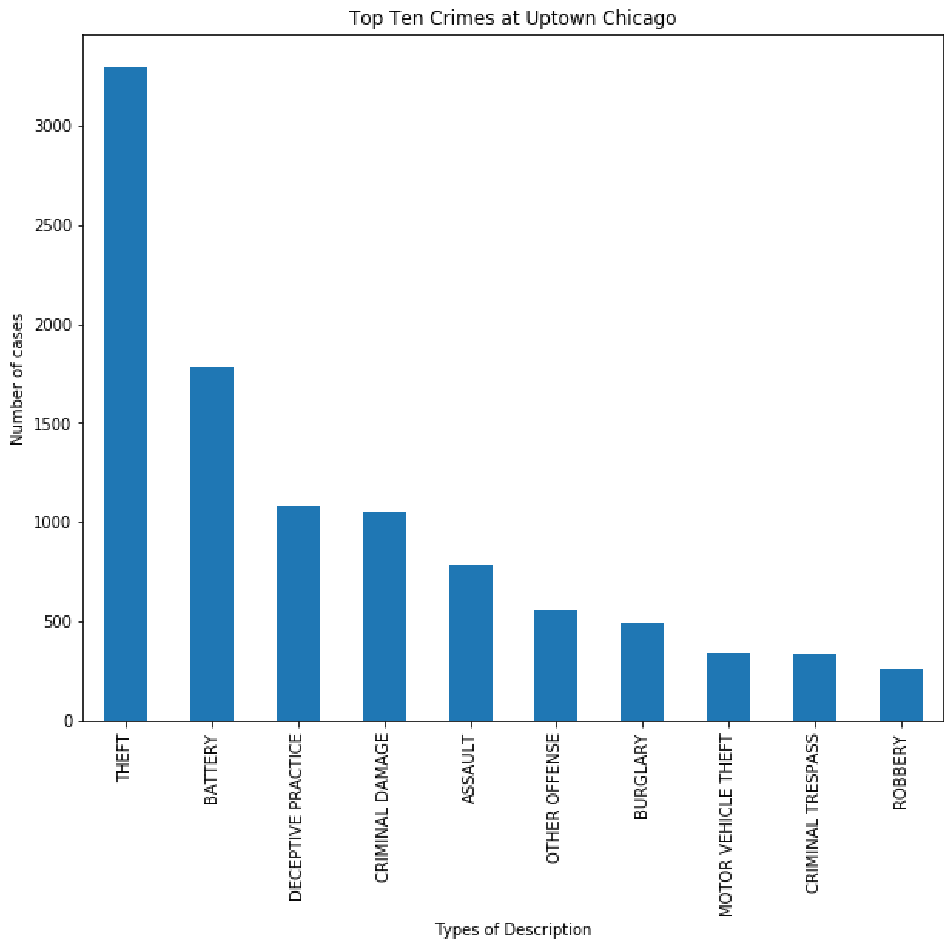
Our nearby_venues function returns 309 venue in all the neighborhood.

```
Uptown

In [149]: Chicago_uptown_venues.shape

Out[149]: (309, 7)
```

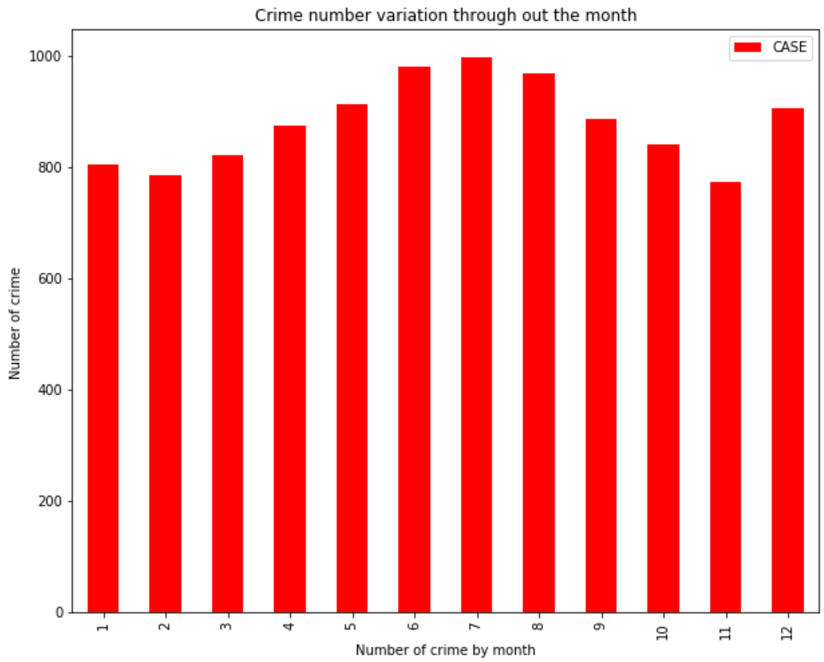
2.3 Data Analysis and Visualizations

Now look at some statistics at uptown_chicago crime data

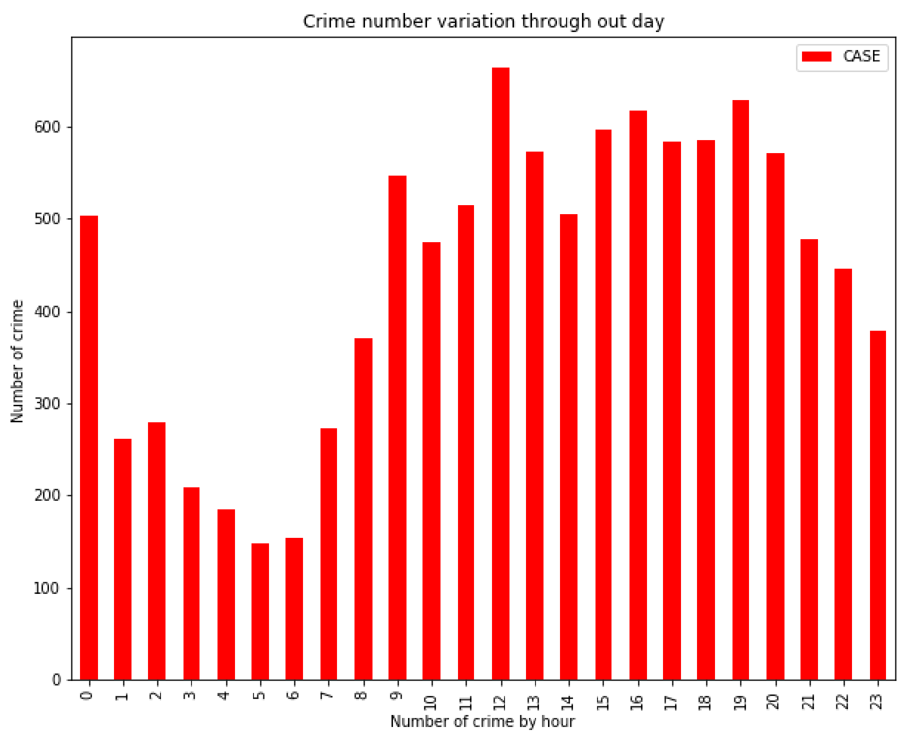
```
In [85]: crime_uptown.DESCRIPTION.value_counts()[:10]

Out[85]: THEFT                  3297
         BATTERY                1779
         DECEPTIVE PRACTICE     1080
         CRIMINAL DAMAGE        1053
         ASSAULT                 789
         OTHER OFFENSE           556
         BURGLARY                493
         MOTOR VEHICLE THEFT     345
         CRIMINAL TRESPASS       333
         ROBBERY                 263
```



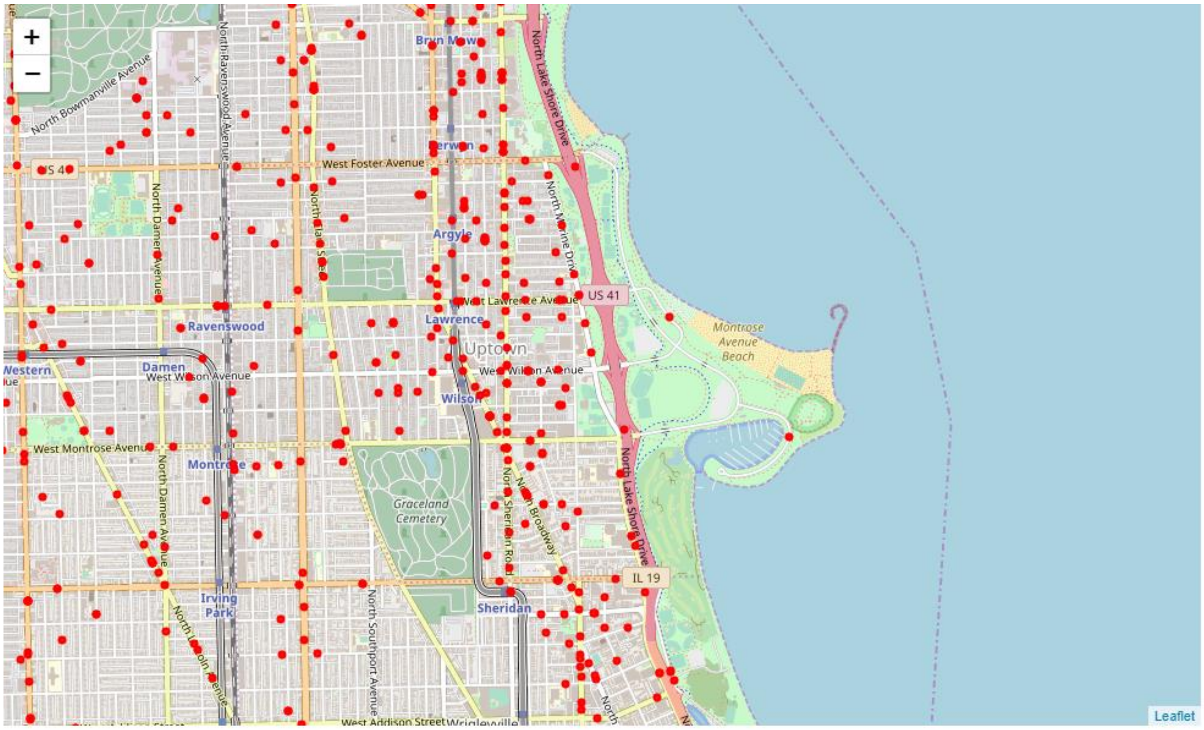
Looking at top 10 crime at uptown Chicago, it is obvious that theft frequency is almost twice and third than respective 2nd and 3rd most frequent crime.



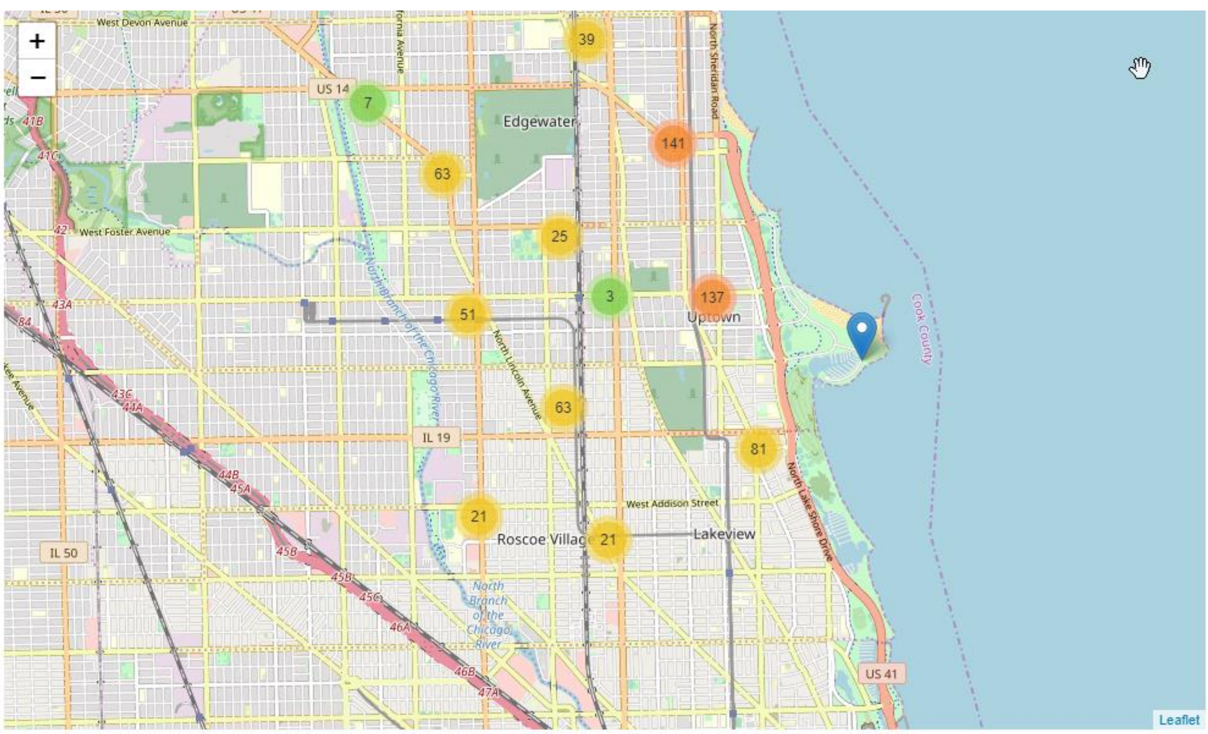
Number of increased in the middle of year especially in June, July, August it is pic.



Number of crime is at its peak during noon. It is higher from 3pm to 9pm.



Crime distributions for 2019 until November15 is superimpose on map created by folium. Then a clustered map is created. It shows center of uptown community has high crime rate.



Methodology

3.1 Data Analysis Description

In this section we analyze our type of venues for each neighborhood with One Hot Encoding method. Machine learning algorithms cannot work with categorical data directly. Categorical data must be converted to numbers. Since we are working with a sequence classification type problem One Hot Encoding will make our categorical data interpretation for machine.

One Hot Encoding for analyze each neighborhood

In [153]:
```
# one hot encoding
Chicago_uptown_onehot = pd.get_dummies(Chicago_uptown_venues[['Venue_Category']], prefix="", prefix_sep="")

# add neighborhood column back to dataframe
Chicago_uptown_onehot['Neighborhood'] = Chicago_uptown_venues['Neighborhood']

# move neighborhood column to the first column
fixed_columns = [Chicago_uptown_onehot.columns[-1]] + list(Chicago_uptown_onehot.columns[:-1])
Chicago_uptown_onehot = Chicago_uptown_onehot[fixed_columns]

Chicago_uptown_onehot.head()
```

Out[153]:

	Neighborhood	American Restaurant	Arcade	Asian Restaurant	Athletics & Sports	Automotive Shop	Bagel Shop	Bakery	Bank	Bar	Basketball Court	Beach	Breakfast Spot	Bubble Tea Shop	Bus Station	Business Service
0	Buena Park	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
1	Buena Park	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	Buena Park	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
3	Buena Park	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	Buena Park	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0

In [154]:
```
uptown_grouped = Chicago_uptown_onehot.groupby('Neighborhood').mean().reset_index()
uptown_grouped.shape
```

Out[154]: (6, 104)

Here is the list of top ten venue in all the 6 neighborhood after one hot encoding

	Neighborhood	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
0	Buena Park	Coffee Shop	Pharmacy	Asian Restaurant	Mexican Restaurant	Bar	Bus Station	Hot Dog Joint	Convenience Store	Flower Shop	Massage Studio
1	Clarendon Park	Park	Convenience Store	Coffee Shop	Beach	American Restaurant	Athletics & Sports	Bus Station	General Entertainment	Mexican Restaurant	Field
2	Margate Park	Vietnamese Restaurant	Chinese Restaurant	Grocery Store	Bus Station	Coffee Shop	American Restaurant	Noodle House	Plaza	Pizza Place	Park
3	New Chinatown	Vietnamese Restaurant	Chinese Restaurant	Grocery Store	Mexican Restaurant	Bubble Tea Shop	Pharmacy	Bakery	Plaza	Market	Bus Station
4	Sheridan Park	Mexican Restaurant	Thai Restaurant	Italian Restaurant	Coffee Shop	Pizza Place	Chinese Restaurant	Bus Station	Liquor Store	Sushi Restaurant	Bar

Now we want to see how similar or dissimilar are all 6 Chicago uptown neighbor hood

3.2What Machine Learning Used

In this problem we want to explore if all the six neighborhood at uptown Chicago is same or they can be grouped in a cluster. Note to mention here our data do not have any labels. Clustering is considered an unsupervised learning method since we don’t have the ground truth to compare the output of the clustering algorithm to the true labels to evaluate its performance. So here we will use **Kmeans algorithm** which is considered as one of the most used clustering algorithms due to its simplicity.

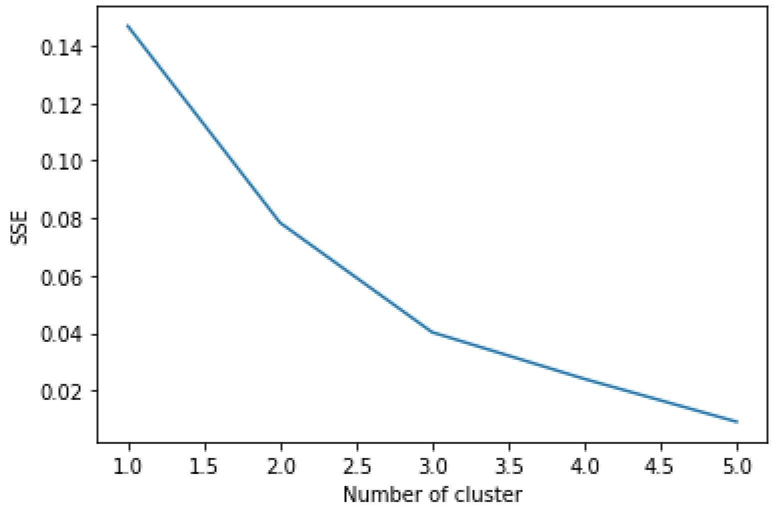
The way kmeans algorithm works is as follows:

1. Specify number of clusters *K*.
2. Initialize centroids by first shuffling the dataset and then randomly selecting *K* data points for the centroids without replacement.
3. Keep iterating until there is no change to the centroids. i.e assignment of data points to clusters isn’t changing. It has steps below:

    - Compute the sum of the squared distance between data points and all centroids.
    - Assign each data point to the closest cluster (centroid).
    - Compute the centroids for the clusters by taking the average of the all data points that belong to each cluster.

3.2 Choosing K

There is a popular method known as **elbow method** which is used to determine the optimal value of K to perform the K-Means Clustering Algorithm. The basic idea behind this method is that it plots the various values of cost with changing *k*. As the value of *K* increases, there will be fewer elements in the cluster. So average distortion will decrease. The lesser number of elements means closer to the centroid. So, the point where this distortion declines the most is the **elbow point**.



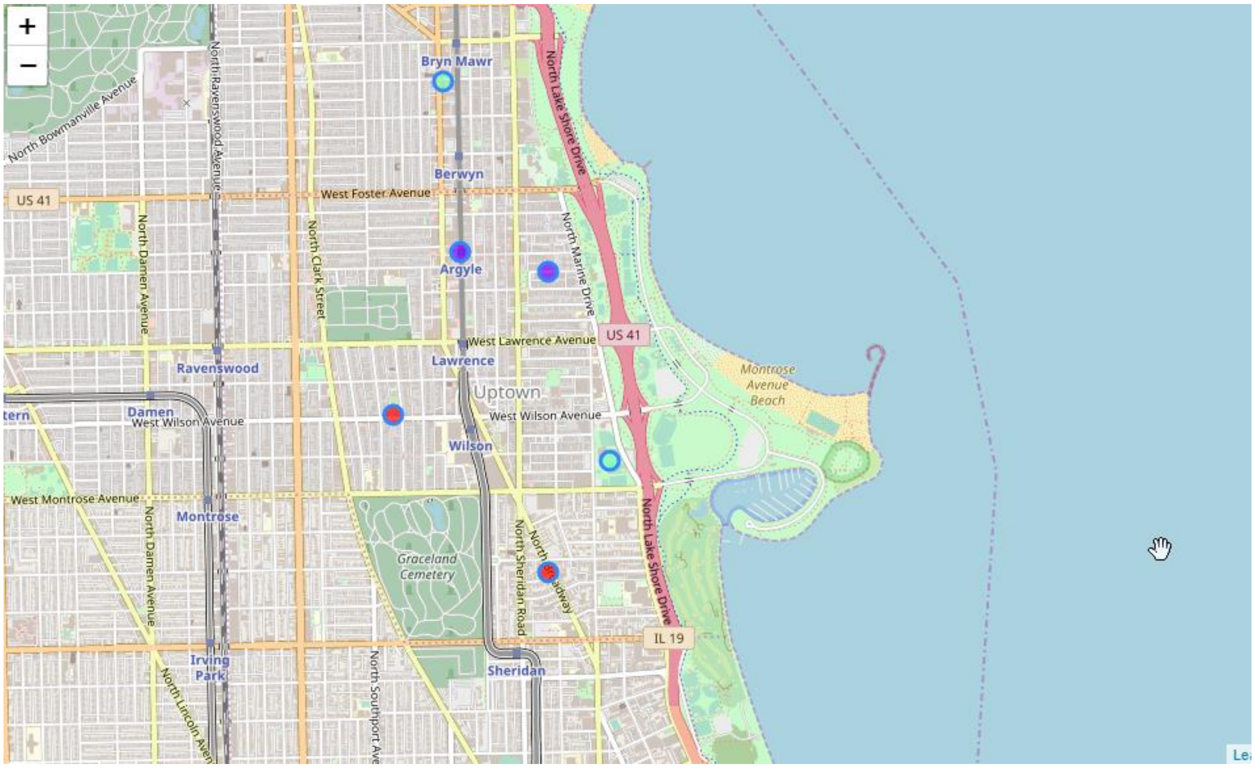
We find our K = 3

Results

4.1 Results

	Neighborhood	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
0	Buena Park	Coffee Shop	Pharmacy	Asian Restaurant	Mexican Restaurant	Bar	Bus Station	Hot Dog Joint	Convenience Store	Flower Shop	Massage Studio
1	Clarendon Park	Park	Convenience Store	Coffee Shop	Beach	American Restaurant	Athletics & Sports	Bus Station	General Entertainment	Mexican Restaurant	Field
2	Margate Park	Vietnamese Restaurant	Chinese Restaurant	Grocery Store	Bus Station	Coffee Shop	American Restaurant	Noodle House	Plaza	Pizza Place	Park
3	New Chinatown	Vietnamese Restaurant	Chinese Restaurant	Grocery Store	Mexican Restaurant	Bubble Tea Shop	Pharmacy	Bakery	Plaza	Market	Bus Station
4	Sheridan Park	Mexican Restaurant	Thai Restaurant	Italian Restaurant	Coffee Shop	Pizza Place	Chinese Restaurant	Bus Station	Liquor Store	Sushi Restaurant	Bar
5	Uptown	Sushi Restaurant	Asian Restaurant	Sandwich Place	Pizza Place	Chinese Restaurant	Vietnamese Restaurant	Bank	Optical Shop	Bus Station	Theater

Here is our neighborhood venue data with Cluster label.



Here neighborhood clusters are on uptown map.

Here is all the clusters

Cluster 0

In [166]:
```
uptown_merged.loc[uptown_merged['Cluster Labels'] == 0, uptown_merged.columns[[0] + list(range(6, uptown_merged.shape[1]))]]
```

Out[166]:

	Neighborhood	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
0	Buena Park	Pharmacy	Asian Restaurant	Mexican Restaurant	Bar	Bus Station	Hot Dog Joint	Convenience Store	Flower Shop	Massage Studio
4	Sheridan Park	Thai Restaurant	Italian Restaurant	Coffee Shop	Pizza Place	Chinese Restaurant	Bus Station	Liquor Store	Sushi Restaurant	Bar

Cluster 1 ¶

```
uptown_merged.loc[uptown_merged['Cluster Labels'] == 1, uptown_merged.columns[[0] + list(range(6, uptown_merged.shape[1]))]]
```

	Neighborhood	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	Margate Park	Chinese Restaurant	Grocery Store	Bus Station	Coffee Shop	American Restaurant	Noodle House	Plaza	Pizza Place	Park
3	New Chinatown	Chinese Restaurant	Grocery Store	Mexican Restaurant	Bubble Tea Shop	Pharmacy	Bakery	Plaza	Market	Bus Station

Cluster 2

```
uptown_merged.loc[uptown_merged['Cluster Labels'] == 2, uptown_merged.columns[[0] + list(range(5, uptown_merged.shape[1]))]]
```

	Neighborhood	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	Clarendon Park	Park	Convenience Store	Coffee Shop	Beach	American Restaurant	Athletics & Sports	Bus Station	General Entertainment	Mexican Restaurant	Field
5	Uptown	Sushi Restaurant	Asian Restaurant	Sandwich Place	Pizza Place	Chinese Restaurant	Vietnamese Restaurant	Bank	Optical Shop	Bus Station	Theater

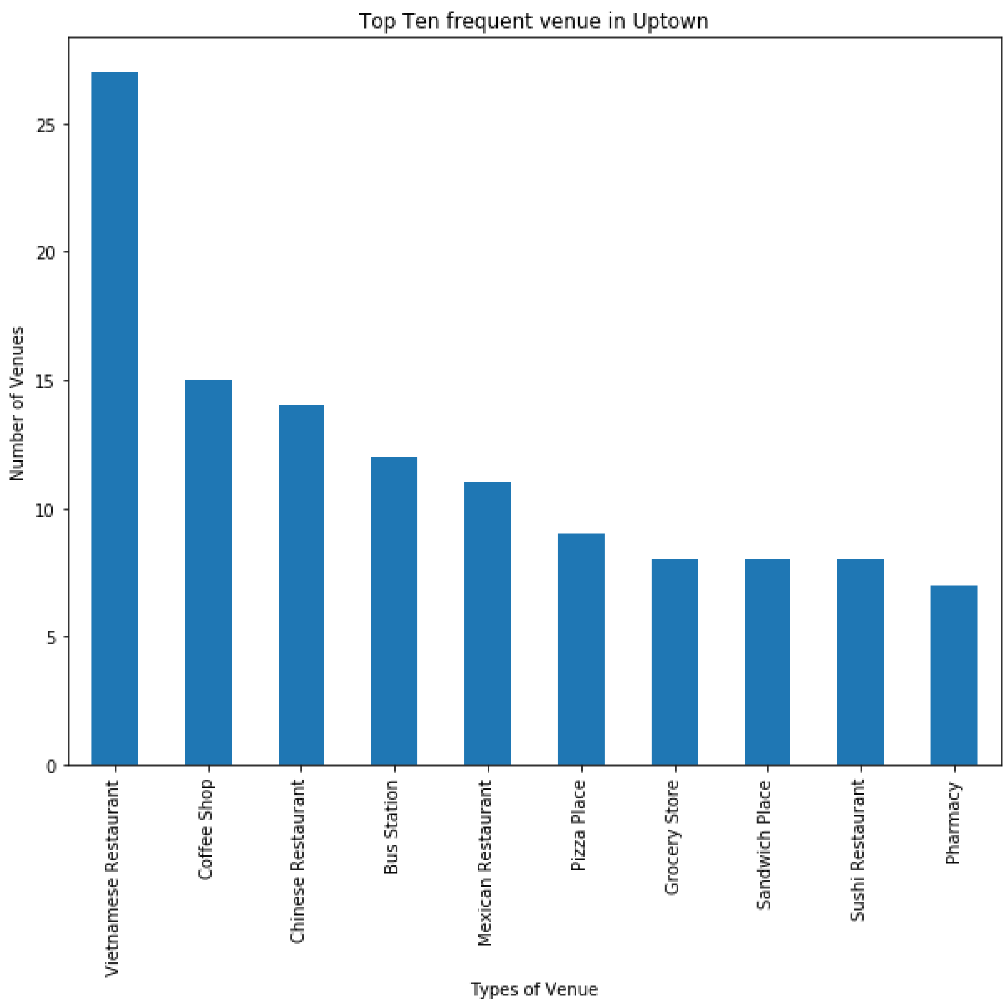
Discussions

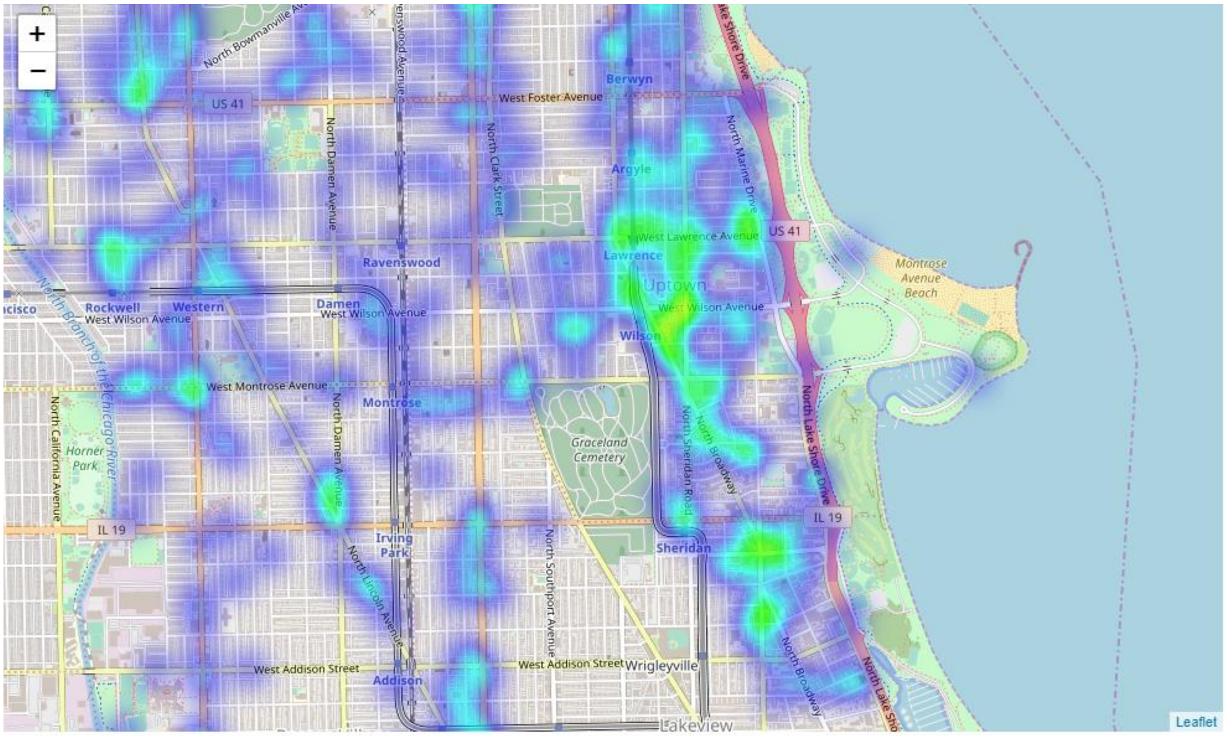
5.1 Observation

Buena Park [shown red in map left side] and Sheridan Park [shown red lower on in map] are clustered together in cluster 0. If anybody wants to have some Mexican food this two neighborhood is the one you are looking for. Both neighborhood has bar and they are among most visited place. These two place has popular Thai, Asian, Chinese, Mexican Italian venue.

If we look carefully at Cluster 1(Shown purple in cluster map) Margate and New Chinatown [It is still marked as its old name Argyle neighborhood on map] are grouped together. If you particularly a fan of Vietnamese or Chinese food this place is perfect for you. This two places grocery store is 3^{nd} most visited place there is strong probability that you can get Asian /tropical fruits or vegetable in those store.

Now in Cluster 2(shown mint in clustered map) consist of Clarendon Park and Uptown. Those two neighborhood are not neighbor to each other. If you are looking for some entertainment or relaxation along with food you could consider either of this two neighborhood. Another things to mention their sushi bar, Vietnamese restaurant and Asian restaurant is popular.





Crime heat map for 2019 Nov crime data

5.2 Recommendations

If anyone wants to explore some Asian ethnic cuisine or Mexican or Italian all 6 uptown neighborhood is worth exploring. Crime heat map shows upper side of uptown community comparatively have less crime. Whereas Clarendon park, New Chinatown and Margate Park has little bit high crime rate.

Theft, battery and assault, deceptive practice are among the top 3 types of crime. It's an indication that this community does not have high volume of gang violence for which Chicago's some areas are notorious for. As a precaution if you ever in this area try to be in a group and pay extra careful about you belongings like cell phone, wallet, watch etc.

For all the 6 neighborhood bus station is always among the top ten popular venue. This is an indication of car parking space shortage. Keeping that in account for exploration public transportation could be a solution to avoid parking hassle.

If you choose any place to eat at Clarendon Park save some of your time to visit park or beach

Conclusion

We explore uptown Chicago community with combination of crime, neighborhood and foursquare API data. This is an excellent way to start working independently on a project and identified potentiality of our self to dig down into the ocean of data and finding hidden gem. Result could be improve by adding venue rating, distance. Crime rate, restaurant number, rating can be compared to nation average with some additional data. Census data could be used in the model to get comparative demographic data with other community