

# Chicago Community Overview using Data Science

Kailas Kundikar

December 25<sup>th</sup>, 2019

## 1. Introduction

### 1.1. Background:

I am from Chicago and resident for 12+ years. Many occupations Chicago visitors has asked me what are top 5 attractions to visits in certain area as they are on short trip. Sometime colleagues who moved to Chicago ask which area is better in terms of venues around or more secure (fewer crimes).

The City of Chicago covers an area of 60,000 hectares and sits 176 meters (578 feet) above sea level on the southwestern shore of Lake Michigan. At 190 km wide and 495 km long, it's the 5th largest body of fresh water in the world. Chicago's extensive parklands, including 3,000 hectares of city parks attract an estimated 86 million visitors annually.

As a multicultural city that thrives on the harmony and diversity of its neighborhoods, Chicago today embodies the values of America's heartland-integrity, hard work and community and reflects the ideals in the social fabric of its **77 distinct neighborhoods**.

Chicago is known all over the United States as a very passionate sports town. Chicago is a leader in reforming public schools, enhancing public safety and security initiatives, providing affordable housing in attractive and economically sound communities, ensuring accessibility for all and fostering, social, economic and environmental sustainability.

A high-level overview of 77 community areas with top 10 spots from each will share good idea on how parks, sport bars and gyms are popular in most of the area to show trends of park & sport friendly city. In addition, crime details will share useful information with residents on which area to choose for safety.

### 1.2. Problem

There are multiple dataset available but no clear visuals and guidelines on how neighborhood around and in terms of safety which community area to pick up for residents.

### 1.3. Area of Interest/Target audience

All Chicago residents and Visitors can benefit from this data analysis and conclusions.

## 2. Data Engineering

### 2.1. Data Sourcing:

Data requirement:

- (1) Chicago Community overview (Cook County).
- (2) Chicago Best top 5 spot not to miss for Tourist in loop area
- (3) Chicago Best spot 10 for residents to hang out per community area.
- (4) Chicago most Safe Community area ranking using Crime data.

Source #1) Wiki page about Chicago Community

Areas [https://en.wikipedia.org/wiki/Community\\_areas\\_in\\_Chicago](https://en.wikipedia.org/wiki/Community_areas_in_Chicago) to get Community Area and neighborhoods

Source #2) Google for latitude and longitude for each community area (RPA process to search and get) in Chicago and prepare excel/csv file to load on IBM Watson.

Source #3) Foursquare API to get venue details per community area

Source #4) Chicago Crime rate data from kaggle.com, for 2012 to 2017

<https://www.kaggle.com/currie32/crimes-in-chicago>

How to Solve?

(1) Go to wiki page about Chicago Community

Areas [https://en.wikipedia.org/wiki/Community\\_areas\\_in\\_Chicago](https://en.wikipedia.org/wiki/Community_areas_in_Chicago) to get Community Area and neighborhoods, it requires complex code to web scraping and combining into single data frame. Using BeautifulSoup4, these multiple tables on wiki page will be fetched and pushed on pandas data frame.

Chicago city portal share longitude and latitude dataset for all areas. Download space details from <https://data.cityofchicago.org/Facilities-Geographic-Boundaries/Boundaries-Community-Areas-current-cauq-8yn6> and connect with above dataset.

(2) Using foursquare developer account get top five best spots around each neighborhood. (Using k-means clustering)

(3) Using foursquare developer count find out places which are top five rated or with most comments. (Using k-means clustering)

(4) Download Chicago Crime rate data from kaggle, for 2012 to 2017: <https://www.kaggle.com/currie32/crimes-in-chicago> based on crime types and count rate the neighbourhoods and share top 5 safe community area to live for. (Using k-means clustering)

## 2.2. Data Cleansing & Integration:

### 2.2.1. Source#1) Wiki page on Chicago community Area and neighborhood

[https://en.wikipedia.org/wiki/Community\\_areas\\_in\\_Chicago](https://en.wikipedia.org/wiki/Community_areas_in_Chicago)

The data is available in multiple tables along with area number and surrounding neighborhoods. It was challenging to pull all in one shot using web scrapping. Using beautiful soup 4 library used to clean tables on html pages and get flat data in rows.

Further Data cleansing required to remove “/n” next line character available on each text fetched and correct the datatype for “Number” field. The multiple neighborhood field has been started from 1<sup>st</sup> charter and separated by “;” for future usage perceptive.

Number	Community area	Neighborhoods
08	Near North Side	<ul style="list-style-type: none"><li>• Cabrini–Green</li><li>• The Gold Coast</li><li>• Goose Island</li><li>• Magnificent Mile</li><li>• Old Town</li><li>• River North</li><li>• River West</li><li>• Streeterville</li></ul>
32	Loop	<ul style="list-style-type: none"><li>• Loop</li><li>• New Eastside</li><li>• South Loop</li><li>• West Loop Gate</li></ul>
33	Near South Side	<ul style="list-style-type: none"><li>• Dearborn Park</li><li>• Printer's Row</li><li>• South Loop</li><li>• Prairie Avenue Historic District</li></ul>

### 2.2.2. Source#2) Global Coordinates for Chicago Community Area

There is no direct data download available for geo coordinates for Chicago Community Area. We have to either manually search each community area on google to get latitude and longitude OR if there would have been more areas, I would have used blue prism to automate the process and dump data into csv file.

e.g. The geographical coordinate of Chicago City are 41.8755616, -87.6244212.

As there are 77 area so manually all area's latitude and longitude looked up and saved in CSV file as below:

Number	Community Area	Latitude	Longitude
8	Near North Side	41.9039	87.6315
32	Loop	41.8786	87.6251
33	Near South Side	41.8608	87.6257
5	North Center	41.9467	87.6883
6	Lake View	41.9398	87.6589
7	Lincoln Park	41.9255	87.6488

Note: the Longitude should be “-ve” number and later in code it has been multiplied by “-1”.

### 2.2.3. Source#3) Chicago community areas and respective venue places.

The venue details and their coordinates are not directly available so using FOURSQUARE developer login account & API provided each areas venue gathered

e.g. URL link for foursquare to get venue details:

```
'https://api.foursquare.com/v2/venues/explore?&client_id=xxxxxxx&client_secret=xxxxxx&v=20191010&ll=41.8786,-87.6251&radius=500&limit=5'
```

The URL take – client id, client secret code, version, geo coordinates of center of area, radius and number of venue to be fetched. It returns JSON file, which further converted into flat data.

### 2.2.4. Source#4) Chicago Crime Dataset

#### Crime in Chicago

The Chicago Police Department's Bureau of Records has tracked crime in Chicago since the beginning of the 20th century. The city's overall crime rate, especially the violent crime rate, is substantially higher than the US average. Chicago was responsible for nearly half of 2016's increase in homicides in the US, though national crime rates stayed near historic lows. As of 2017, Chicago's homicide rate is significantly higher when compared to the larger American cities of New York and Los Angeles, but lower when compared to smaller American cities. The reason for the violence, which is localized, to some areas of the city, including change in police tactics or increase in gang rivalry, remain unclear.

Data cleansing, removed the rows where few feature column values are NaN or blank. Below columns were found with missing values and their % compared to total rows available.

Your selected dataframe has 23 columns.  
There are 10 columns that have missing values.

	Missing Values	% of Total Values
X Coordinate	37083	2.5
Y Coordinate	37083	2.5
Latitude	37083	2.5
Longitude	37083	2.5
Location	37083	2.5
Location Description	1658	0.1
Community Area	40	0.0
Ward	14	0.0
Case Number	1	0.0
District	1	0.0

Why is this dataset more important now?

- Chicago Crime Numbers have been in the media for all the wrong reasons.
- President widely used these numbers during his presidential campaign
- He still talks about it.
- These numbers are affecting the overall numbers of the country.
- Did homicides rates really getting worse in Chicago in 2016?
  - Crimes have decreased in Chicago in 2016 compared to 2012.
  - There is clear indication in the time series plot that crime numbers increase somewhere during middle of the year mostly during summer months and drops during end/starting of the year mostly during winter months.
  - Why are the number of arrests so less compared to number of crimes?
  - It would be interesting to look at how the crimes are distributed by months.
- The Chicago Crime dataset contains a summary of the reported crimes occurred in the City of Chicago from 2012 to 2017.
- Dataset has been obtained from the Chicago Police Department's CLEAR (Citizen Law Enforcement Analysis and Reporting) system.
- Dataset contains the following columns:
  - **ID:** Unique identifier for the record.
  - **Case Number:** The Chicago Police Department RD Number (Records Division Number), which is unique to the incident.
  - **Date:** Date when the incident occurred.
  - **Block:** address where the incident occurred
  - **IUCR:** The Illinois Uniform Crime Reporting code.
  - **Primary Type:** The primary description of the IUCR code.
  - **Description:** The secondary description of the IUCR code, a subcategory of the primary description.
  - **Location Description:** Description of the location where the incident occurred.
  - **Arrest:** Indicates whether an arrest was made.

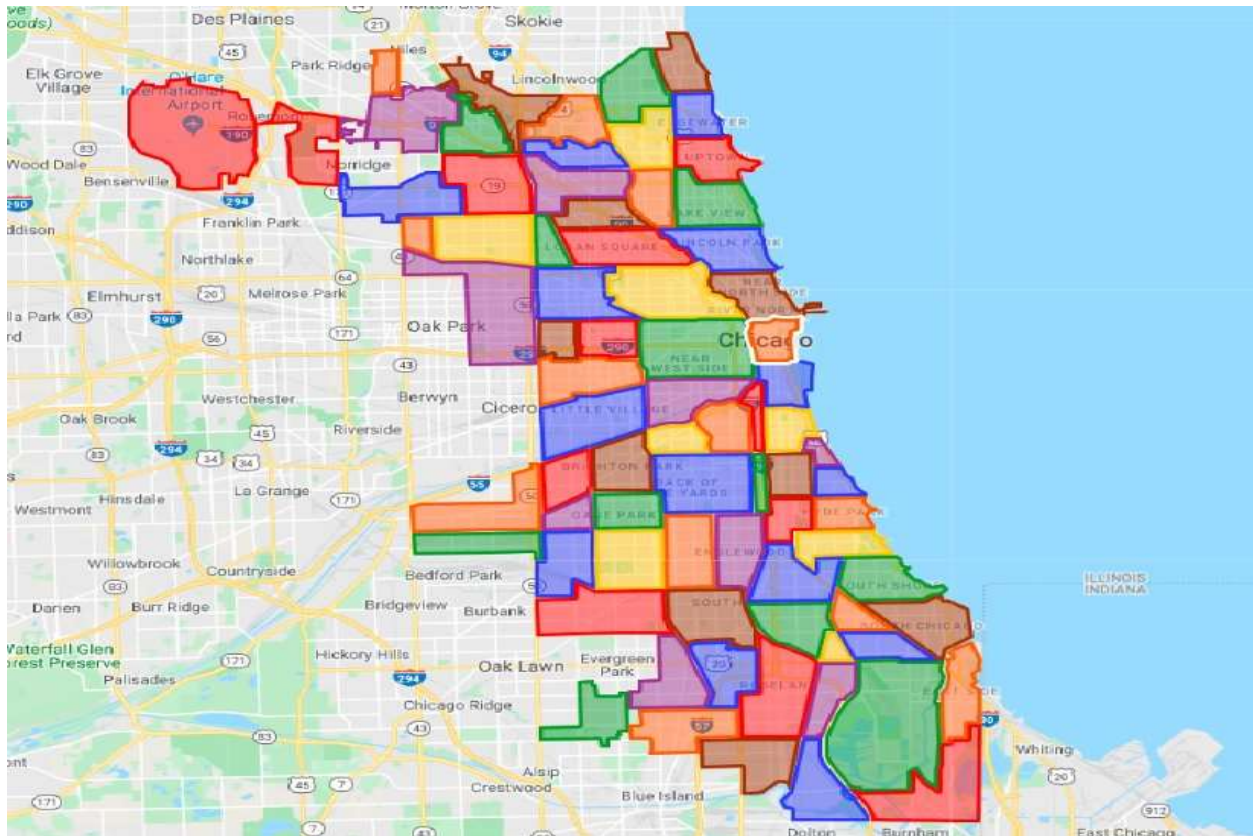
- **Domestic:** Indicates whether the incident was domestic-related as defined by the Illinois Domestic Violence Act.
- **Beat:** Indicates the beat where the incident occurred. A beat is the smallest police geographic area – each beat has a dedicated police beat car.
- **District:** Indicates the police district where the incident occurred.
- **Ward:** The ward (City Council district) where the incident occurred.
- **Community Area:** Indicates the community area where the incident occurred. Chicago has 77 community areas.
- **FBI Code:** Indicates the crime classification as outlined in the FBI's National Incident-Based Reporting System (NIBRS).
- **X Coordinate:** The x coordinate of the location where the incident occurred in State Plane Illinois East NAD 1983 projection.
- **Y Coordinate:** The y coordinate of the location where the incident occurred in State Plane Illinois East NAD 1983 projection.
- **Year:** Year the incident occurred.
- **Updated On:** Date and time the record was last updated.
- **Latitude:** 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:** 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:** 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.

## 2.3. Data Profiling:

### Data Objectives:

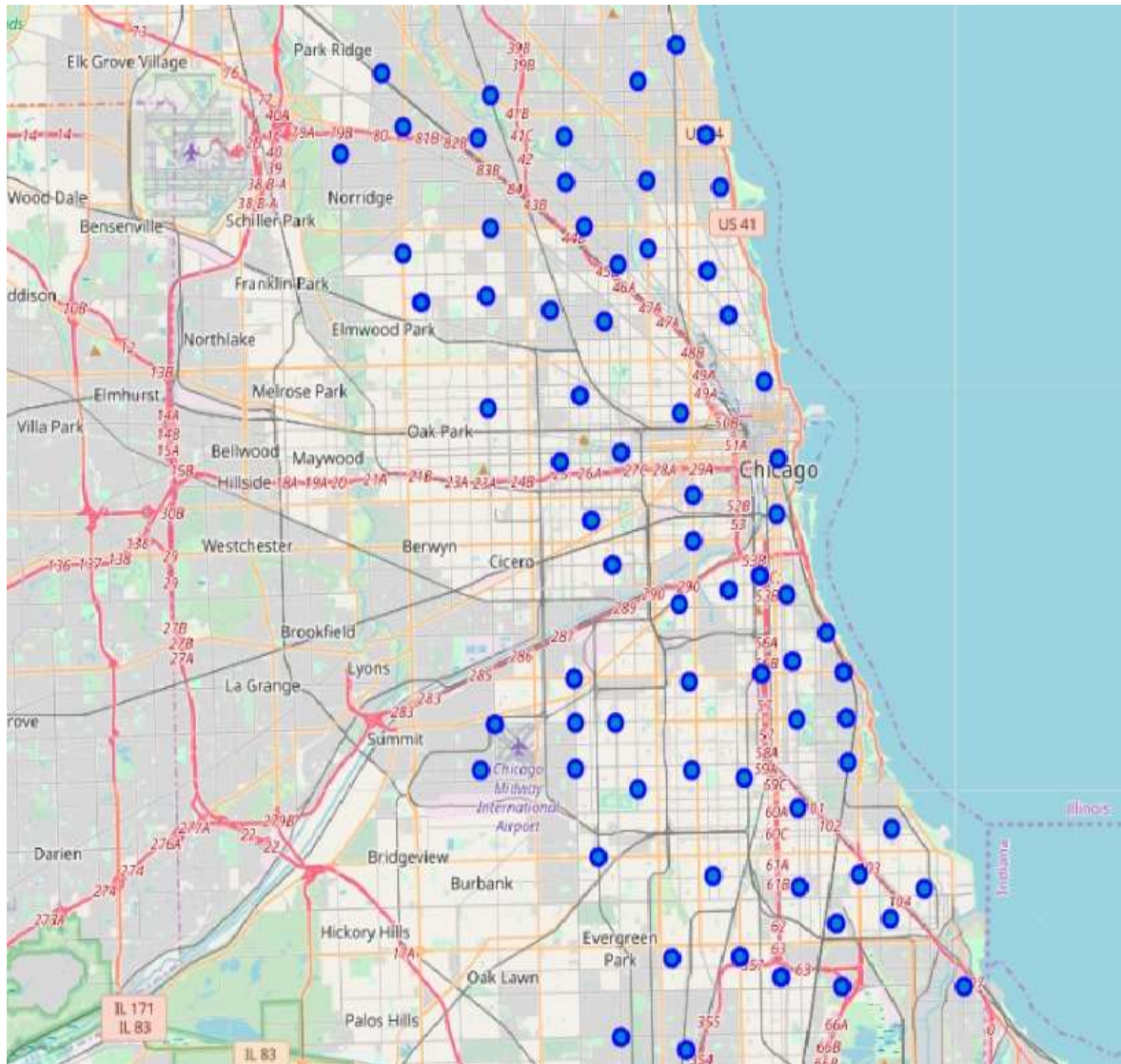
#### 2.3.1. (1) Chicago Community overview (Cook County).

77 community areas along with boundaries.



For each community area, center location identified and plotted using folium. The center location used to find venues around 500 m of radius.





### 2.3.2. (2) Top 5 attractions in LOOP / downtown area for Tourist.

Chicago downtown or sometime referred, as loop is always been attraction of many visitors and residents. It has beautiful skyline and many museums around. The top 5 spots that close to loop are as below:

	name	categories	lat	lng
0	The Art Institute of Chicago	Art Museum	41.879610	-87.623552
1	Symphony Center (Chicago Symphony Orchestra)	Concert Hall	41.879275	-87.624680
2	Chicago Architecture Foundation	Museum	41.878556	-87.624550
3	Thome Miniature Rooms	Museum	41.879532	-87.623680
4	The South Garden	Garden	41.878626	-87.624207



Using folium map their exact location is as below:



### 2.3.3. (3) Top 10 venues for residents per community area.

There are total 77 community areas and few of them might of skipped in API calls due to safety reason or few might not famous venues in their 500m radius. Using foursquare API venue around center location of community fetched.

Raw/flat data from API:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Near North Side	41.9039	-87.6315	Chicago Q	41.903599	-87.630277	BBQ Joint
1	Near North Side	41.9039	-87.6315	Sparrow	41.903277	-87.628972	Cocktail Bar
2	Near North Side	41.9039	-87.6315	3 Arts Club Cafe	41.905800	-87.630470	Café
3	Near North Side	41.9039	-87.6315	Division Street Farmers Market	41.903908	-87.630015	Farmers Market
4	Near North Side	41.9039	-87.6315	BIGGS, Chicago's Cigar Shop and Smoking Lounge	41.903201	-87.630254	Lounge

Using one hot encoding and transposing to get one row per area and venues as columns:

	Neighborhood	ATM	African Restaurant	American Restaurant	Arepa Restaurant	Art Gallery	Art Museum	Asian Restaurant	Athletics & Sports	BBQ Joint	...	Trail	Train Station	Vegetarian / Vegan Restaurant	Video Game Store	Video Store	Vietnamese Restaurant	Weight Loss Center	Wine Shop	Wings Joint	Yoga Studio
0	Albany Park	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	Archer Heights	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
2	Armour Square	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	Ashburn	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	Auburn Gresham	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Converting one hot encoding and picking up top 10 places and putting them order:

	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	Albany Park	Korean Restaurant	Dive Bar	Discount Store	Mobile Phone Shop	Seafood Restaurant	Fried Chicken Joint	Pizza Place	Coffee Shop	Pet Store	Department Store
1	Archer Heights	Mexican Restaurant	Bank	Grocery Store	Video Store	Seafood Restaurant	Nightclub	Rental Car Location	Bakery	Bar	Dive Bar
2	Armour Square	Chinese Restaurant	Grocery Store	Sandwich Place	Sports Bar	Cosmetics Shop	Gas Station	Hot Dog Joint	Italian Restaurant	Fast Food Restaurant	Farmers Market
3	Ashburn	Cosmetics Shop	Mexican Restaurant	Light Rail Station	Fried Chicken Joint	Eastern European Restaurant	Filipino Restaurant	Fast Food Restaurant	Farmers Market	Eye Doctor	Ethiopian Restaurant
4	Auburn Gresham	Discount Store	Pool	Park	Basketball Court	Donut Shop	Filipino Restaurant	Fast Food Restaurant	Farmers Market	Eye Doctor	Ethiopian Restaurant
5	Austin	Breakfast Spot	Discount Store	Pizza Place	BBQ Joint	Food	Café	Thrift / Vintage Store	Cosmetics Shop	Dessert Shop	Diner
6	Avalon Park	Burger Joint	Grocery Store	Diner	Fast Food Restaurant	Cajun / Creole Restaurant	Sandwich Place	Boutique	Pizza Place	Home Service	Dessert Shop
7	Avondale	Brewery	Beer Store	Burger Joint	German Restaurant	Pub	Park	Gaming Cafe	Restaurant	Italian Restaurant	Fried Chicken Joint

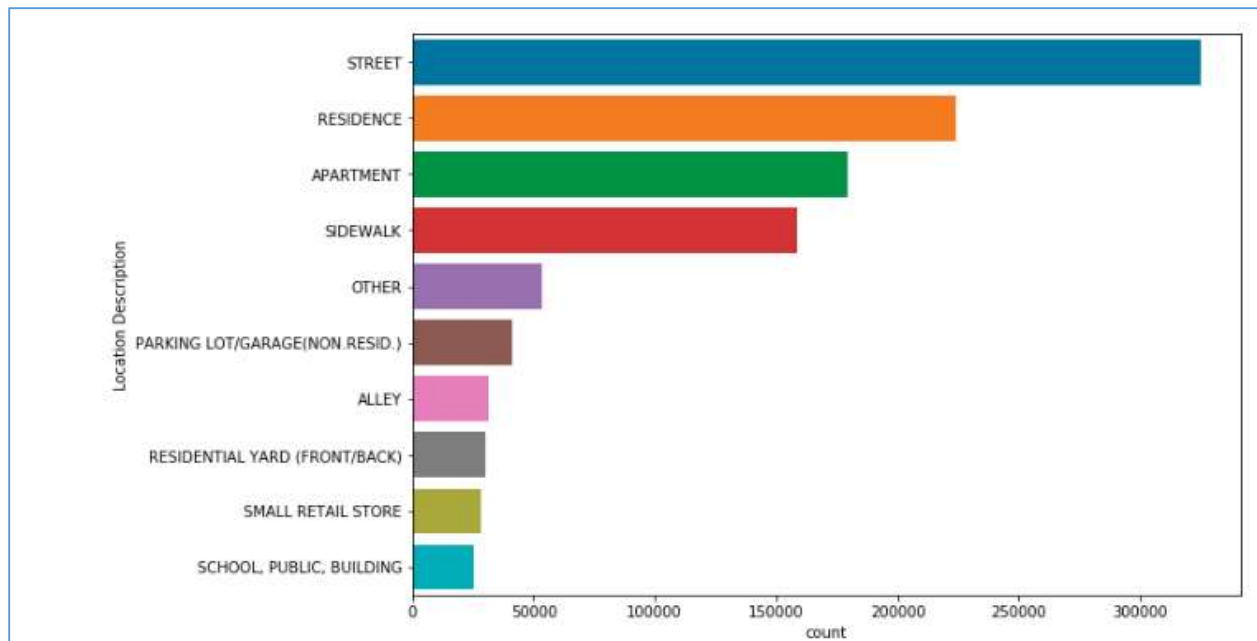
#### 2.3.4. (4) Rank Chicago community areas for safe living based on Crime data.

Chicago dataset has multiple columns related to crime. To get feature columns, data profiling done to get familiar with each attribute.

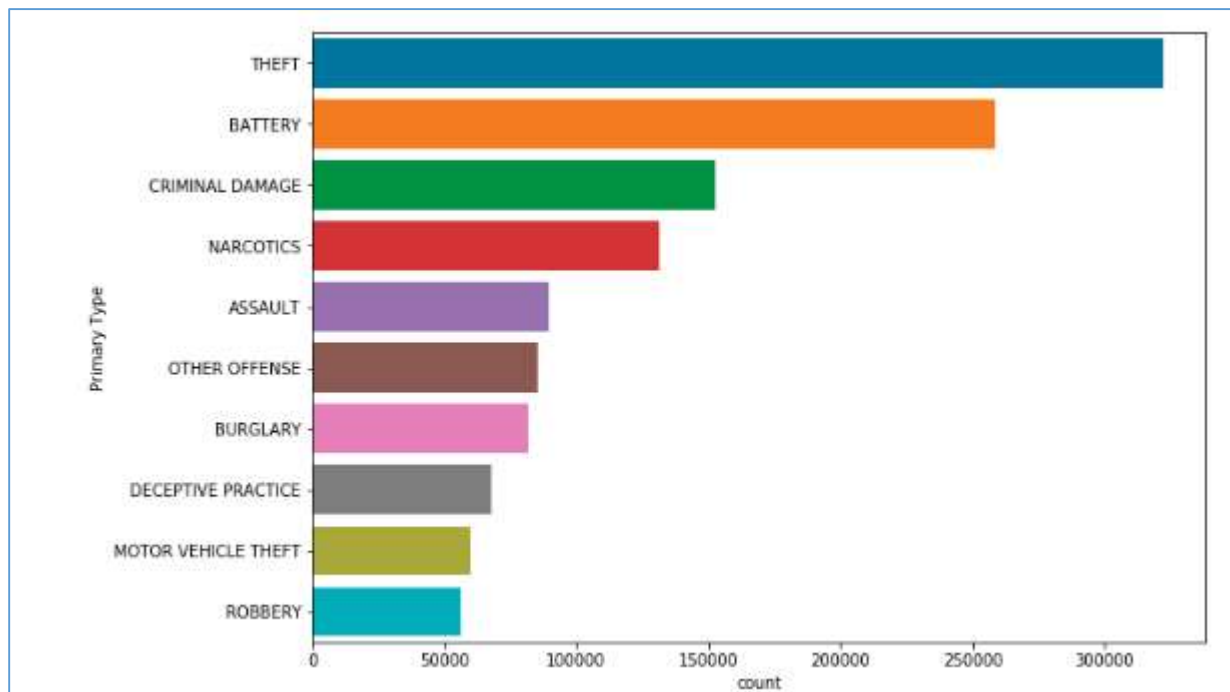
Unnamed: 0	ID	Case Number	Date	Block	IUCR	Primary Type	Description	Location Description	Arrest	Domestic	Beat	District	Ward	Community Area	FBI Code	X Coordinate	Y Coordinate	Year	Updated On	La	
0	3	10508693	H2250496	05/03/2016 11:40:00 PM	013XX S SAWYER AVE	0486	BATTERY	DOMESTIC BATTERY SIMPLE	APARTMENT	True	True	1022	10.0	24.0	29.0	08B	1154907.0	1893681.0	2016	05/10/2016 03:56:50 PM	41.8
1	89	10508695	H2250409	05/03/2016 09:40:00 PM	061XX S DREXEL AVE	0486	BATTERY	DOMESTIC BATTERY SIMPLE	RESIDENCE	False	True	313	3.0	20.0	42.0	08B	1183066.0	1864330.0	2016	05/10/2016 03:56:50 PM	41.7
2	197	10508697	H2250503	05/03/2016 11:31:00 PM	053XX W CHICAGO AVE	0470	PUBLIC PEACE VIOLATION	RECKLESS CONDUCT	STREET	False	False	1524	15.0	37.0	25.0	24	1140789.0	1904819.0	2016	05/10/2016 03:56:50 PM	41.8
3	673	10508698	H2250424	05/03/2016 10:10:00 PM	049XX W FULTON ST	0460	BATTERY	SIMPLE	SIDEWALK	False	False	1532	15.0	28.0	25.0	08B	1143223.0	1901475.0	2016	05/10/2016 03:56:50 PM	41.8
4	911	10508699	H2250455	05/03/2016 10:00:00 PM	003XX N LOTUS AVE	0820	THEFT	\$500 AND UNDER	RESIDENCE	False	True	1523	15.0	28.0	25.0	06	1139890.0	1901675.0	2016	05/10/2016 03:56:50 PM	41.8

As per data profiling Crime date, Arrest (after crime), crime location and Primary type of crime attributes found important to further drill down.

Base on 'location description' column found that most of crimes were on street.

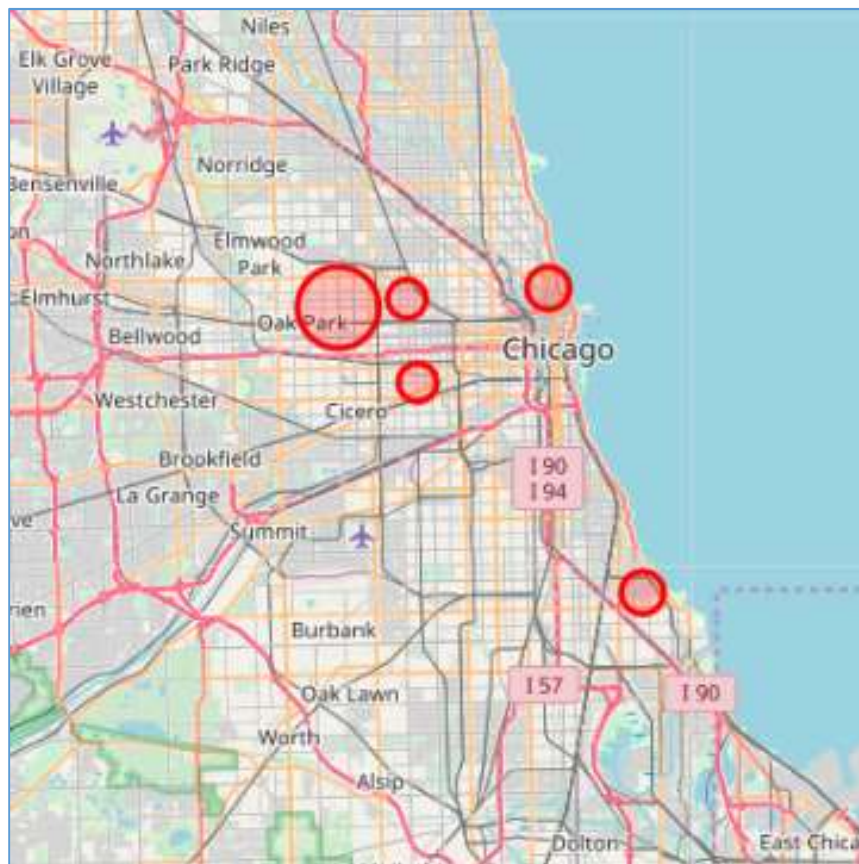
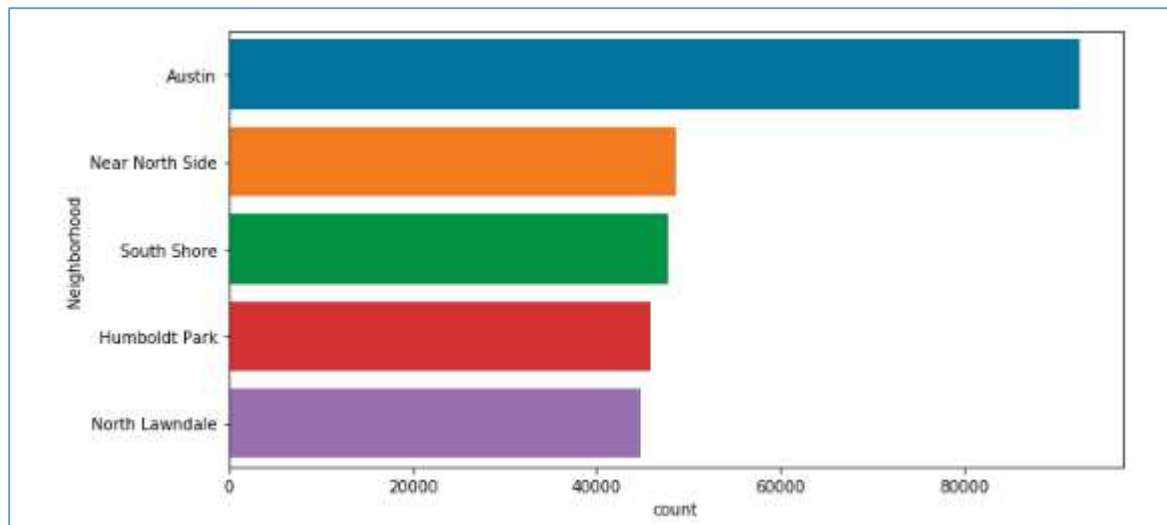


Based on data profiling of 'primary type' column, it is found that most of crime are of 'theft' type:



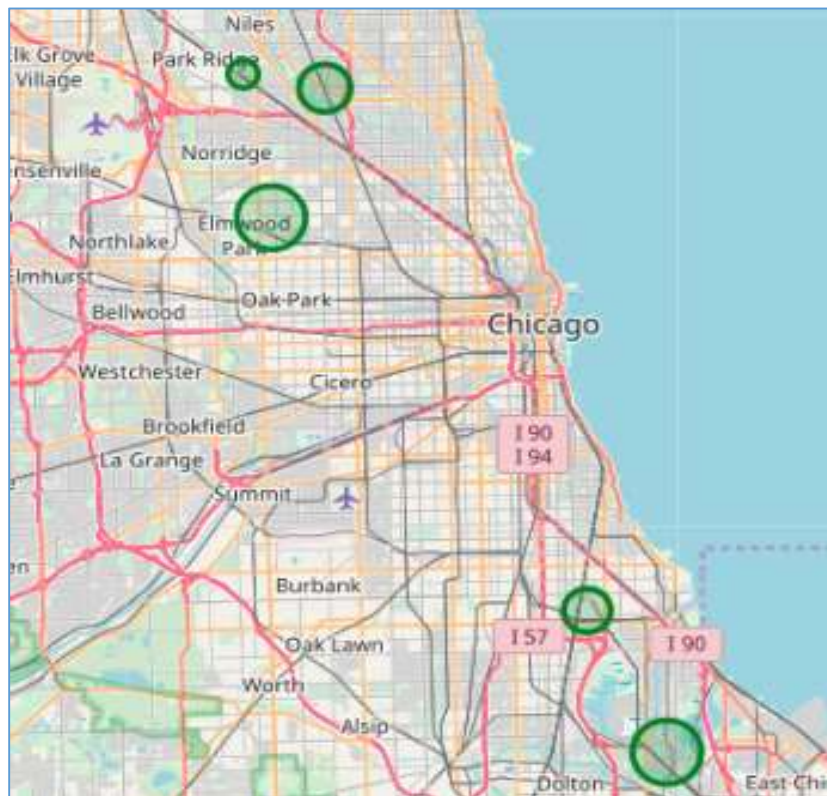
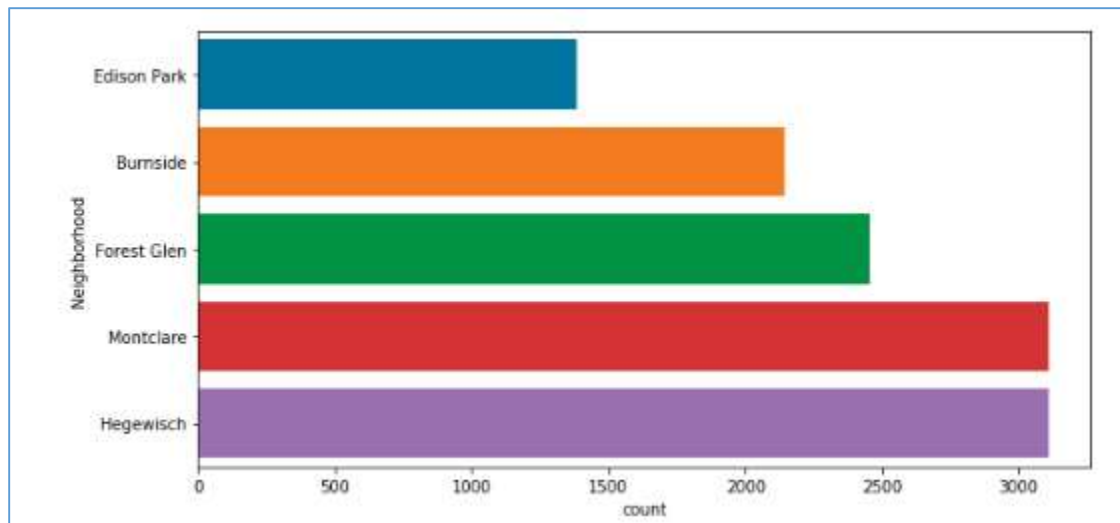
Crime data linked with Chicago community area dataset to get community area name and then grouping data points to get view of crime count in each area.

Top 5 **not safe area** based on crime count are as below:



Top 5 **safe** area based on crime count are as below:





### 3. Machine learning

#### 3.1. Classification model: K means clustering

Cluster1: Due to lot of Pizza places as Chicago downtown and uptown area is famous for Pizza.

Neighborhood	Cluster Labels	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
North Center	0	Pizza Place	Gym / Fitness Center	Bar	Coffee Shop	Sandwich Place	Bowling Alley	Football Stadium	Liquor Store	Grocery Store	Garden Center
Lake View	0	Performing Arts Venue	Pizza Place	Vegetarian / Vegan Restaurant	Martial Arts Dojo	Japanese Restaurant	Taco Place	Bakery	Fish & Chips Shop	Coffee Shop	Eastern European Restaurant
Edison Park	0	Italian Restaurant	Pizza Place	Mexican Restaurant	Pub	Deli / Bodega	Hot Dog Joint	French Restaurant	Bakery	Bar	Donut Shop
Albany Park	0	Korean Restaurant	Pizza Place	Fried Chicken Joint	Pet Store	Coffee Shop	Mobile Phone Shop	Seafood Restaurant	Dive Bar	Discount Store	Ethiopian Restaurant
Dunning	0	Thai Restaurant	Hot Dog Joint	Asian Restaurant	Italian Restaurant	Tattoo Parlor	Ice Cream Shop	Pizza Place	Gourmet Shop	Deli / Bodega	Department Store
Montclare	0	Pizza Place	Bar	Convenience Store	Fast Food Restaurant	Flea Market	Cycle Studio	Eastern European Restaurant	Cosmetics Shop	Filipino Restaurant	Cupcake Shop
Near West Side	0	Italian Restaurant	Frozen Yogurt Shop	Bar	Breakfast Spot	Supplement Shop	Mexican Restaurant	Ice Cream Shop	Sandwich Place	Farmers Market	Eye Doctor
Bridgeport	0	Bar	Coffee Shop	Breakfast Spot	Pizza Place	Mediterranean Restaurant	Art Gallery	Korean Restaurant	Café	Fast Food Restaurant	Farmers Market
Garfield Ridge	0	Pizza Place	Liquor Store	Chinese Restaurant	Spa	Supermarket	Breakfast Spot	Boutique	Convenience Store	Pharmacy	Department Store
South Chicago	0	Pizza Place	Supermarket	Train Station	Ice Cream Shop	Italian Restaurant	Discount Store	Farmers Market	Eye Doctor	Ethiopian Restaurant	Eastern European Restaurant
East Side	0	Pizza Place	Deli / Bodega	Shipping Store	Italian Restaurant	Rental Car Location	Bakery	Cycle Studio	Department Store	Currency Exchange	Dessert Shop
Beverly	0	Pizza Place	Cosmetics Shop	Grocery Store	Breakfast Spot	Deli / Bodega	Burger Joint	Boutique	Caribbean Restaurant	Italian Restaurant	Eastern European Restaurant

## Cluster2: fast food place or coffee places

Neighborhood	Cluster Labels	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
Near North Side	1	Lounge	Cocktail Bar	Farmers Market	Café	Restaurant	BBQ Joint	Furniture / Home Store	Cycle Studio	Gym	Pizza Place
Loop	1	Museum	Garden	Donut Shop	Concert Hall	Arepa Restaurant	Art Museum	Asian Restaurant	Hot Dog Joint	Sushi Restaurant	Fast Food Restaurant
Near South Side	1	Tapas Restaurant	Lounge	Park	Optical Shop	Cocktail Bar	New American Restaurant	Sushi Restaurant	Café	Italian Restaurant	Bar
Lincoln Park	1	Music Venue	Pizza Place	Pub	Israeli Restaurant	Indonesian Restaurant	Theater	Coffee Shop	Cupcake Shop	Gym / Fitness Center	Eye Doctor
Avondale	1	Restaurant	Pub	Burger Joint	Beer Store	Italian Restaurant	Fried Chicken Joint	Gaming Cafe	Brewery	Park	German Restaurant
Logan Square	1	Latin American Restaurant	Gym / Fitness Center	Bar	Park	Coffee Shop	Board Shop	Bistro	Bookstore	Historic Site	Gas Station
West Ridge	1	Indian Restaurant	Pakistani Restaurant	Grocery Store	Fruit & Vegetable Store	Park	Farmers Market	Eye Doctor	Ethiopian Restaurant	Eastern European Restaurant	Donut Shop
Uptown	1	Coffee Shop	Deli / Bodega	Lounge	Ethiopian Restaurant	Sushi Restaurant	Tattoo Parlor	Thai Restaurant	Fast Food Restaurant	Vietnamese Restaurant	General Entertainment
Lincoln Square	1	Gourmet Shop	Yoga Studio	Sporting Goods Shop	Filipino Restaurant	Pedestrian Plaza	Café	Butcher	Bookstore	Gym	Auto Workshop
Forest Glen	1	Sandwich Place	Coffee Shop	Train Station	Salon / Barbershop	Diner	Trail	Ice Cream Shop	Vietnamese Restaurant	Park	Currency Exchange
Edgewater	1	Asian Restaurant	Gym	Tanning Salon	Shipping Store	Vietnamese Restaurant	American Restaurant	Hobby Shop	Italian Restaurant	Bakery	Filipino Restaurant
Portage Park	1	Fast Food Restaurant	Greek Restaurant	Pizza Place	Gym	Farmers Market	Sports Bar	Coffee Shop	Ice Cream Shop	Park	Bike Shop

## Cluster3:Mexican restaurants



Neighborhood	Cluster Labels	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
Rogers Park	2	Mexican Restaurant	Park	Pet Store	Deli / Bodega	Pizza Place	Sandwich Place	Yoga Studio	Discount Store	Eye Doctor	Ethiopian Restaurant
Jefferson Park	2	American Restaurant	Pizza Place	Deli / Bodega	Burger Joint	Hot Dog Joint	Mexican Restaurant	Spanish Restaurant	Bar	Yoga Studio	Eastern European Restaurant
Belmont Cragin	2	Mexican Restaurant	Pharmacy	Thrift / Vintage Store	Mobile Phone Shop	Discount Store	Bakery	Nightclub	Grocery Store	American Restaurant	Deli / Bodega
Hermosa	2	Mexican Restaurant	Grocery Store	Greek Restaurant	Latin American Restaurant	Park	Seafood Restaurant	BBQ Joint	Bus Station	Donut Shop	Farmers Market
South Lawndale	2	Mexican Restaurant	Taco Place	Grocery Store	Restaurant	Sculpture Garden	Dive Bar	Farmers Market	Eye Doctor	Ethiopian Restaurant	Eastern European Restaurant
Lower West Side	2	Mexican Restaurant	Grocery Store	Music Venue	Supermarket	Brewery	Bar	Bakery	Pizza Place	Dessert Shop	Diner
Archer Heights	2	Mexican Restaurant	Bakery	Video Store	Grocery Store	Seafood Restaurant	Nightclub	Rental Car Location	Bank	Bar	Discount Store
Brighton Park	2	Mexican Restaurant	Bakery	Video Store	Grocery Store	Seafood Restaurant	Nightclub	Rental Car Location	Bank	Bar	Discount Store
McKinley Park	2	Diner	Coffee Shop	Big Box Store	Seafood Restaurant	Mexican Restaurant	Discount Store	Grocery Store	Video Game Store	Video Store	Cycle Studio
New City	2	Pharmacy	Bank	Dessert Shop	Discount Store	Mexican Restaurant	Fried Chicken Joint	Grocery Store	Video Store	Gas Station	Diner

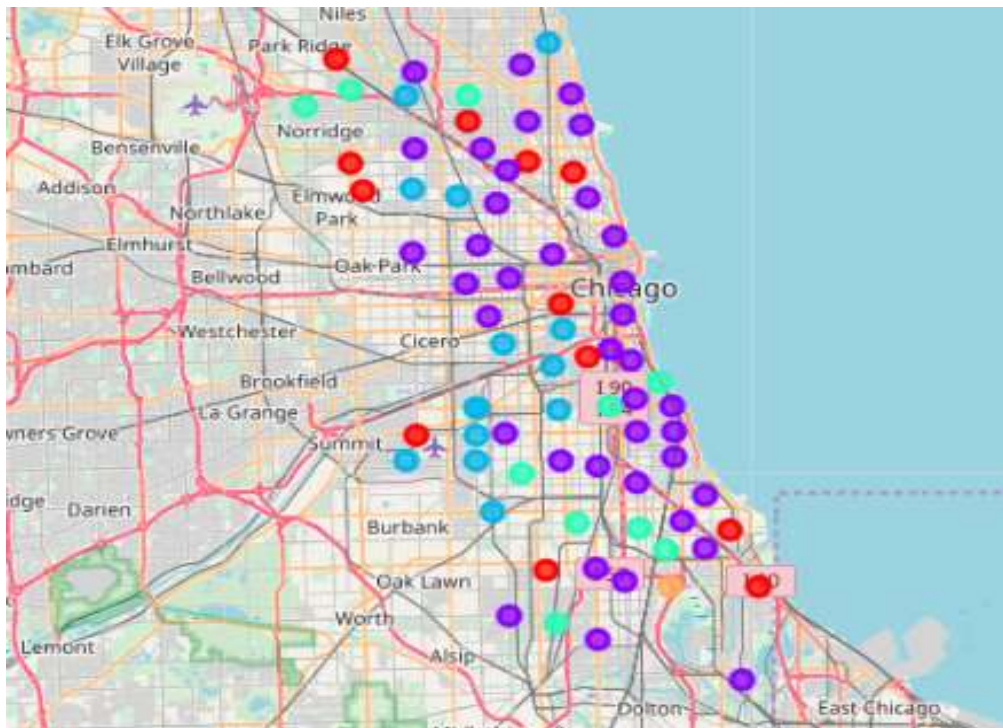
#### Cluster4:

Neighborhood	Cluster Labels	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
Norwood Park	3	Speakeasy	Clothing Store	Park	Road	Yoga Studio	Discount Store	Farmers Market	Eye Doctor	Ethiopian Restaurant	Eastern European Restaurant
North Park	3	Gymnastics Gym	Department Store	Park	Dive Bar	Filipino Restaurant	Fast Food Restaurant	Farmers Market	Eye Doctor	Ethiopian Restaurant	Eastern European Restaurant
O'Hare	3	Convenience Store	Home Service	Eastern European Restaurant	Dive Bar	Park	Insurance Office	American Restaurant	Deli / Bodega	Department Store	Cycle Studio
Oakland	3	Beach	Park	Trail	Track	Yoga Studio	Dive Bar	Fast Food Restaurant	Farmers Market	Eye Doctor	Ethiopian Restaurant
Fuller Park	3	Bus Station	Park	Train Station	Food	Fast Food Restaurant	Department Store	Baseball Field	Yoga Studio	Farmers Market	Eye Doctor
Chicago Lawn	3	Fish & Chips Shop	Intersection	American Restaurant	Fast Food Restaurant	Park	Filipino Restaurant	Farmers Market	Eye Doctor	Ethiopian Restaurant	Eastern European Restaurant
Chatham	3	Park	Boutique	Fast Food Restaurant	Fried Chicken Joint	Discount Store	Lounge	Donut Shop	Yoga Studio	Ethiopian Restaurant	Filipino Restaurant
Burnside	3	Intersection	Food	Home Service	Park	Discount Store	Fast Food Restaurant	Farmers Market	Eye Doctor	Ethiopian Restaurant	Eastern European Restaurant
Auburn Gresham	3	Pool	Park	Basketball Court	Discount Store	Yoga Studio	Dive Bar	Fast Food Restaurant	Farmers Market	Eye Doctor	Ethiopian Restaurant
Morgan Park	3	Construction & Landscaping	Currency Exchange	Park	Train Station	Dive Bar	Filipino Restaurant	Fast Food Restaurant	Farmers Market	Eye Doctor	Ethiopian Restaurant

Cluster5: Looks like odd man, where no restaurants or park in vicinity

Neighborhood	Cluster Labels	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
Pullman	4	ATM	Convenience Store	Fish & Chips Shop	Filipino Restaurant	Fast Food Restaurant	Farmers Market	Eye Doctor	Ethiopian Restaurant	Eastern European Restaurant	Donut Shop

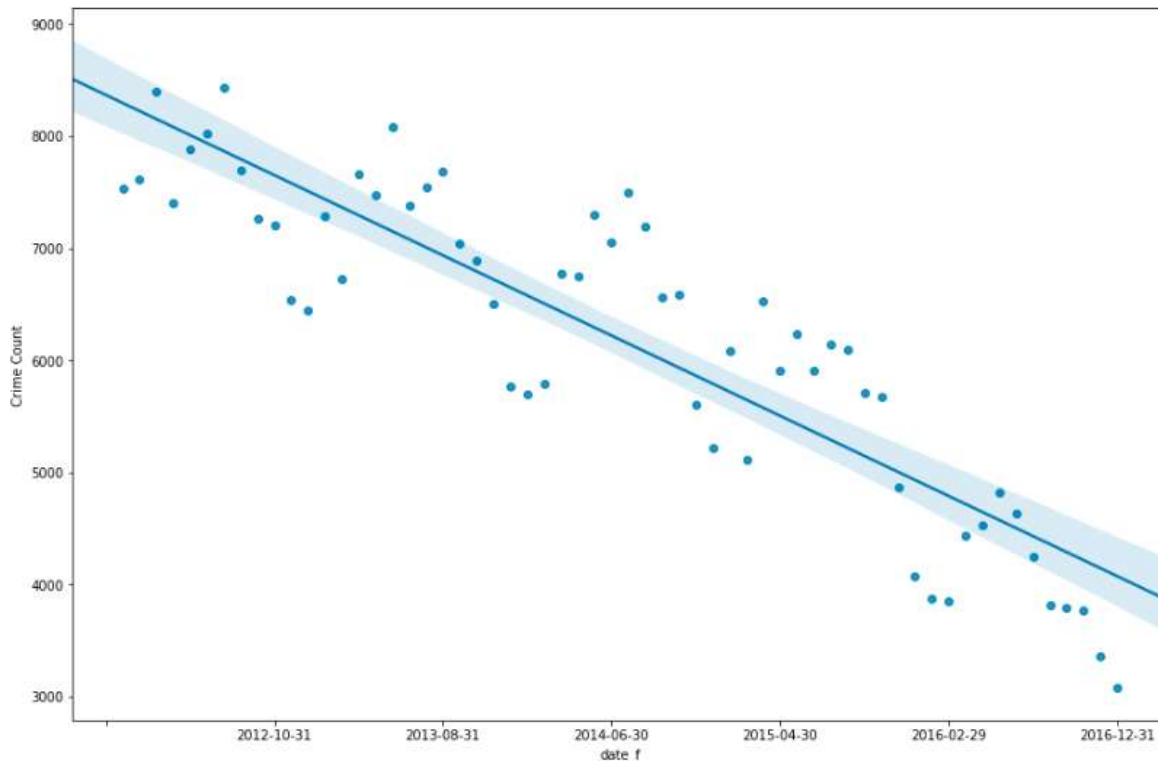
Map with all 5 clusters with color code:



Cluster1 (highlighted **Purple**) Restaurant and fast food place like sandwich or coffee shops all around.  
Cluster2 (highlighted **Red**) famous Pizza restaurants around.  
Cluster3 (highlighted **light green**), Parking spots around.  
Cluster4 (highlighted **light blue**), Mexican restaurants around  
Cluster5 (highlighted **orange**) odd man out with mixed venue.

### 3.2.Predictive model: Linear regression

Linear regression plot:



$$Y_{\text{Predict}} = b + mx$$

Where,

$Y_{\text{Predict}}$ : Number of Crimes that may occur in future

$m$ : slop

$b$ : Coefficient

$x$  : Future date.

**Formula:**

(Y) Predict Crime Count = **8366.88926554** ( $b$ ) - **71.56085024** ( $m$ ) \* ( $x$ ) Future Date Value

## 4. Conclusion

- (1) Overview of 77 community area in Chicago and plot on map
- (2) In loop area top 5 venues found
- (3) In each community area top 10 venues found and areas are clustered as per venue liking.
- (4) Chicago crime data analysis and top 5 safe and top 5 not safe community areas.  
Additional findings are
  - 4.1) Crimes rate has be decreased.
  - 4.2) Crimes rate, on given year, is more during summer time.
  - 4.3) Predictive analysis (linear regression) for future dates.