

Finding the best neighborhood in Chicago

Devi Selvakumar

1.Introduction

1.1 Background

Chicago is the Nation's third largest city and it is truly a melting pot of diverse culture. It sits on the southwestern shore of Lake Michigan, giving it the name "Windy city". The summer time alongside of the lake and the bike paths for miles are breath taking. Chicago's skyline, prairie style architecture, varied culinary delights and abundant green spaces makes it mix big city style with midwestern soul. From 37 movable bridges across the Chicago river to famous art museums, it is made up of diverse neighborhoods each with its distinct vibe and it is home to 10 million residents. It is also known for its crime rate. Personally, I am curious about how we can analyze the available data sets to find the best neighborhood in Chicago.

1.2 Problem

There are several factors to be considered in order to move into a good neighborhood. As Chicago is well known for crime incidents, safety is the foremost critical criteria that should be considered along with close proximity to shopping and restaurants. Even though there are other factors which contributes to the selection of the best neighborhood to move in, for the sake of analysis and availability of data sets, only the above factors are considered. The objective of this project is to explore the Chicago neighborhoods for crime incidents and those who have easy access to stores and restaurants and to find the best neighborhood.

1.3 Stakeholders

Anybody who is interested in finding out the best neighborhood in Chicago or trying to move in to the best Chicago neighborhood with less crime incidents and quick access to restaurants and shopping centers would be potential stakeholders for this project.

2. Data Collection

2.1 Data source

There are two data sets to be used for this project.

Chicago crime data set:

Chicago's current crime data is acquired from the City of Chicago data portal <https://data.cityofchicago.org/api/views/qzdf-xmn8/rows.csv> . The best part of this data set is: it has the crime records from Jan.1st 2020 to Dec.5th 2020, which is very rare. The CSV file lists the crime type, action taken, the district/ward/neighborhood it happened along with other data.

Chicago Neighborhood data set:

There is no direct data set available which lists the community name along with the area code and the map coordinates. Upon thorough search, I came across a text file which has the needed data. This text file parsed the Chicago Neighborhood map and listed the map coordinates of all the points within the neighborhood boundaries along with the name. To locate the community's name with latitude / longitudes, the map of Chicago neighborhoods is processed and the various points in the neighborhoods is listed as a text file in:

https://raw.githubusercontent.com/jkgiesler/parse-chicago-neighborhoods/master/community_to_gps.txt

The text file had 1048576 rows and 4 columns. In total, there are 77 neighborhoods taken into account.

Community Area	Community Name	Latitude	Longitude
1	Rogers Park	42.003801	-87.657651
1	Rogers Park	42.002439	-87.657809
1	Rogers Park	41.99839	-87.657676
1	Rogers Park	42.009069	-87.661341
1	Rogers Park	42.00568	-87.660129
1	Rogers Park	42.00567	-87.661205
1	Rogers Park	42.005401	-87.661198
1	Rogers Park	42.006879	-87.66174
1	Rogers Park	42.001313	-87.662222
1	Rogers Park	42.001551	-87.663739
1	Rogers Park	41.998281	-87.664583
1	Rogers Park	42.009012	-87.664783
1	Rogers Park	42.00806	-87.665325
1	Rogers Park	42.005597	-87.666079
1	Rogers Park	42.007759	-87.666653
1	Rogers Park	42.019371	-87.667178
1	Rogers Park	41.998205	-87.66811
1	Rogers Park	42.01924	-87.668418

3. Methodology

3.1 Feature selection

3.1.1 Chicago crime data set

The data set contains 1048576 rows and 24 columns containing the crime data details from Jan.1,2020 to Dec.5,2020.

ID	Case No.	Date	Block	IUCR	Primary	Descrip	Locatio	Arrest	Domest	Beat	District	Ward	Commu	FBI Cod	X Coord	Y Coord	Year	Upd
12240781	JD453981	1/1/2020 0:00	040XX W C	1154	DECEPTIVE FINANCIAL	MOVIE HC	FALSE	FALSE	1114	11	28	26	11				2020	12/1
12220321	JD430436	1/1/2020 0:00	091XX S D	1752	OFFENSE I AGGRAVA	RESIDENCI	FALSE	TRUE	413	4	8	47	17				2020	11/1
12191982	JD397743	1/1/2020 0:00	116XX S E	1153	DECEPTIVE FINANCIAL	APARTMEI	FALSE	FALSE	522	5	34	53	11	1175385	1827727		2020	10/1
12186705	JD391621	1/1/2020 0:00	058XX W C	1153	DECEPTIVE FINANCIAL	OTHER (SF	FALSE	FALSE	1511	15	29	25	11	1137369	1904748		2020	10/1
12181533	JD384685	1/1/2020 0:00	044XX S G	1754	OFFENSE I AGGRAVA	APARTMEI	FALSE	TRUE	222	2	4	39	2	1184334	1875715		2020	10/1
12168471	ID370494	1/1/2020 0:00	049XX N K	281	CRIMINAL	NON-AGG APARTMFI	FAI SF	FAI SF	1712	17	39	14	2	1145802	1932546		2020	9/2

Even though we can look into all crime data, the details of which district/ward/block, case #, ID #, FBI code, update/ arrest status are irrelevant to our current scope of the project. It would have been the best scenario, if it has the community name along with the code. We don't need another data set to match the community code with the community name. As it has only the community area code where the crime has happened, I decided to take into account only 3 features viz. the crime type, date and the community area code. The date is converted from Date Time format to Date format. And, not all community area code has the latitude/longitudes listed in the table. To avoid error, the map coordinates of the neighborhoods are taken from the Neighborhood data set.

	Community Area	Date	Primary Type
0	8	2020-04-10	DECEPTIVE PRACTICE
1	47	2020-12-03	OTHER OFFENSE
2	66	2020-12-03	WEAPONS VIOLATION
3	46	2020-12-03	THEFT
4	71	2020-12-03	CRIMINAL DAMAGE
5	30	2020-11-27	DECEPTIVE PRACTICE
6	61	2020-11-30	THEFT
7	29	2020-12-03	ROBBERY
8	32	2020-12-03	THEFT
9	68	2020-12-03	CRIMINAL DAMAGE
10	24	2020-12-03	MOTOR VEHICLE THEFT
11	48	2020-11-25	DECEPTIVE PRACTICE
12	48	2020-12-01	DECEPTIVE PRACTICE

3.1.2 Chicago Community data

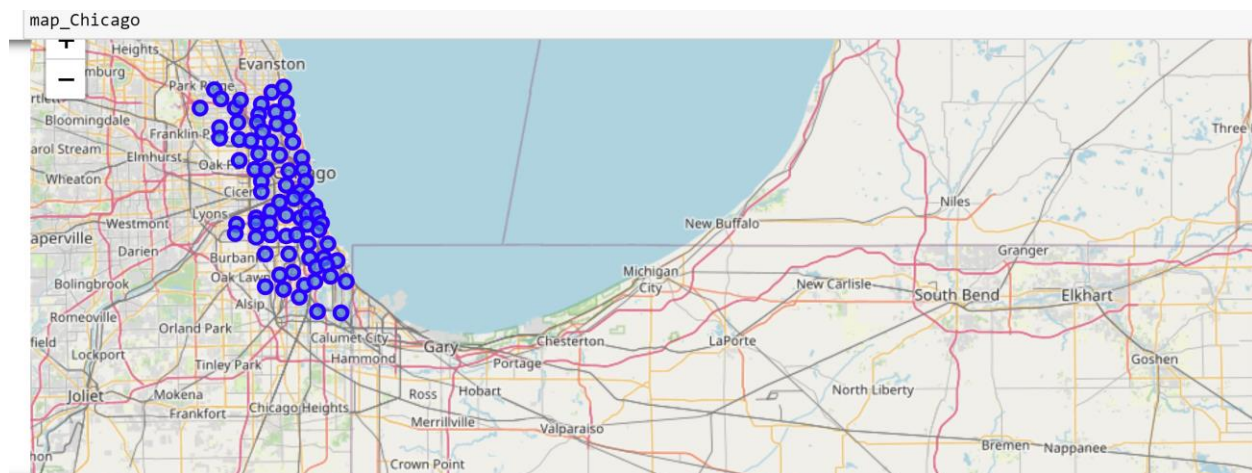
There is no direct data set available which lists the community name along with the code and the map coordinates. Upon thorough search, I came across a text file which has the needed data. This text file parsed from the Chicago Neighborhood map with the map coordinates of all the points within the neighborhood boundaries along with the name. The original text file had 1048576 rows and 4 columns. I have decided to take the average of latitude/longitude for each community area to use for the purpose of the project and the final data frame contains 77 rows and 4 columns.

	Community Area	Community Name	Latitude	Longitude
0	1	Rogers Park	42.009120	-87.668648
1	2	West Ridge	41.999316	-87.692394
2	3	Uptown	41.966222	-87.658792
3	4	Lincoln Square	41.968844	-87.685397
4	5	North Center	41.950503	-87.681029
...
72	73	Washington Heights	41.717396	-87.650298
73	74	Mount Greenwood	41.695023	-87.706722
74	75	Morgan Park	41.691034	-87.667514
75	76	O'Hare	41.975927	-87.843830
76	77	Edgewater	41.984299	-87.663061

3.2 Data visualization

Using Folium, the map is produced to show how the neighborhoods are spread across Chicago.

Chicago neighborhood map



Using Foursquare API, the venues near the neighborhood are explored and the data is passed on to the unsupervised learning model. Using One-hot Encoding method, the distinct venue categories are made into unique columns. If the neighborhood venue belongs to that category, it gets a value 1 for that column and 0 if not. The final output is a sparse matrix with 1's & 0's

representing whether the neighborhood venue belong to that category or not. The data frame is grouped by the community name is shown with the average values shown for each category.

Neighborhood	ATM	Accessories Store	Afghan Restaurant	African Restaurant	Airport	American Restaurant	Antique Shop	Arcade	Arepa Restaurant	...	Vegetarian / Vegan Restaurant	Video Game Store	Video Store	Vietnamese Restaurant
1	0.0	0.0	0.000	0.0	0.0	0.062500	0.000000	0.0000	0.0	...	0.000000	0.0	0.0	0.000000
2	0.0	0.0	0.025	0.0	0.0	0.025000	0.000000	0.0000	0.0	...	0.000000	0.0	0.0	0.000000
3	0.0	0.0	0.000	0.0	0.0	0.000000	0.000000	0.0125	0.0	...	0.012500	0.0	0.0	0.012500
4	0.0	0.0	0.000	0.0	0.0	0.024390	0.000000	0.0000	0.0	...	0.000000	0.0	0.0	0.000000
5	0.0	0.0	0.000	0.0	0.0	0.030303	0.030303	0.0000	0.0	...	0.000000	0.0	0.0	0.000000
...
73	0.0	0.0	0.000	0.0	0.0	0.000000	0.000000	0.0000	0.0	...	0.000000	0.0	0.0	0.000000
74	0.0	0.0	0.000	0.0	0.0	0.000000	0.000000	0.0000	0.0	...	0.000000	0.0	0.0	0.000000
75	0.0	0.0	0.000	0.0	0.0	0.000000	0.000000	0.0000	0.0	...	0.000000	0.0	0.0	0.000000
76	0.0	0.0	0.000	0.0	0.0	0.166667	0.000000	0.0000	0.0	...	0.000000	0.0	0.0	0.000000
77	0.0	0.0	0.000	0.0	0.0	0.000000	0.020833	0.0000	0.0	...	0.020833	0.0	0.0	0.020833

For each neighborhood, the top 10 most common venues are listed in the data frame:

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	Mexican Restaurant	Theater	American Restaurant	Bar	Donut Shop	Bakery	Pizza Place	Diner	Greek Restaurant	Breakfast Spot
2	Indian Restaurant	Pakistani Restaurant	Grocery Store	Donut Shop	Burmese Restaurant	Fruit & Vegetable Store	Market	Chinese Restaurant	Fast Food Restaurant	Bar
3	Coffee Shop	Pizza Place	Bar	Mexican Restaurant	Chinese Restaurant	Sandwich Place	Diner	Bus Station	Sushi Restaurant	Convenience Store
4	Bar	Sandwich Place	Café	Thai Restaurant	Sushi Restaurant	American Restaurant	Art Gallery	Pizza Place	Mobile Phone Shop	Chinese Restaurant
5	Pub	Mobile Phone Shop	Bar	Pizza Place	Coffee Shop	Boutique	Beer Garden	Bank	German Restaurant	Sushi Restaurant

The way the best neighborhood will be sought is by finding the following scores for each neighborhood:

1. Crime score →
 = 0 if the total number of crimes in that neighborhood > the mean of total number of crimes in all neighborhoods
 = 1 if the total number of crimes in that neighborhood < the mean of total number of crimes in all neighborhoods.
2. Shopping score →
 = 0 if the total number of shopping places in that neighborhood < the mean of total number of shopping centers in all neighborhoods
 = 1 if the total number of shopping places in that neighborhood >= the mean of total number of shopping centers in all neighborhoods

3. Eatery score →

= 0 if the total number of restaurant/eating places in that neighborhood < the mean of total number of eatery places in all neighborhoods

= 1 if the total number of restaurant/eating places in that neighborhood ≥ the mean of total number of eatery places in all neighborhoods

4. Total score = Crime score + Shopping score + Eatery score

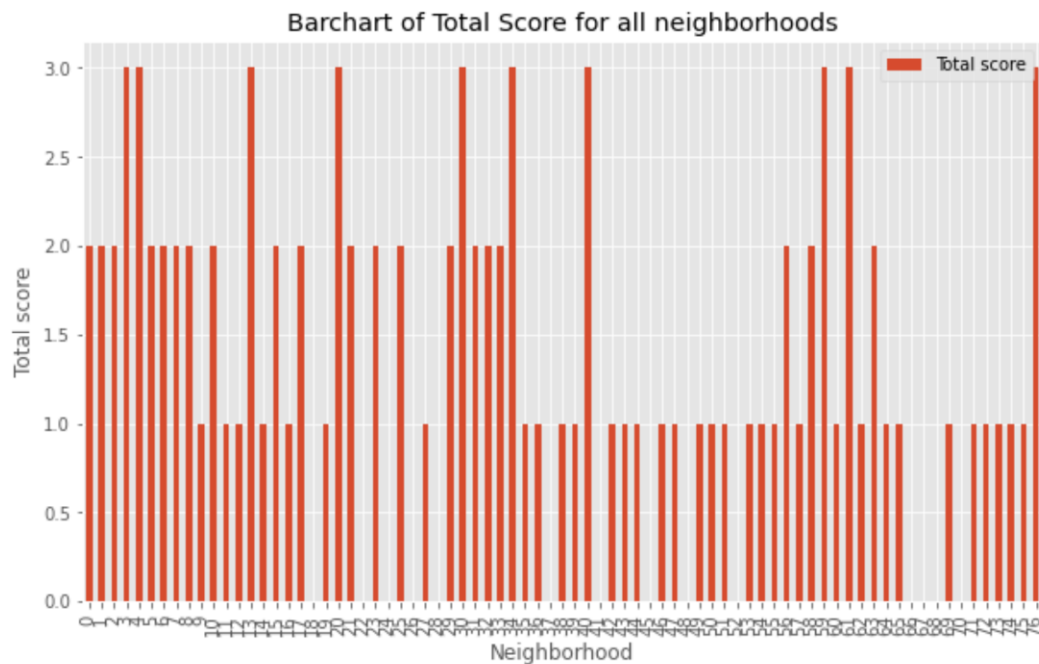
So the best score for each neighborhood is determined as:

- If the total score = 3, then it is the best choice, as
 - the # of crimes in that neighborhood is less than the average
 - The # of shopping centers is more than the average of all shopping places included
 - The # of eatery places are more than the average of all restaurants and other fast-food places.
- If the total score = 2 then it is moderate, 2 of the scores are good and one of them may be because of:
 - Either the # of crimes in that neighborhood is more than the average
 - Or the # of shopping places is less than the average number of shopping centers
 - Or the # of eatery places may be less than the average.
- If the total score = 1, then it is the least choice, as only one of the scores is looking good and other scores are bad.
- If the total score = 0, then it should be avoided, as the total # of crimes is more than the average and the shopping and restaurants are less than the average available in Chicago.

The scores are listed for each community area.

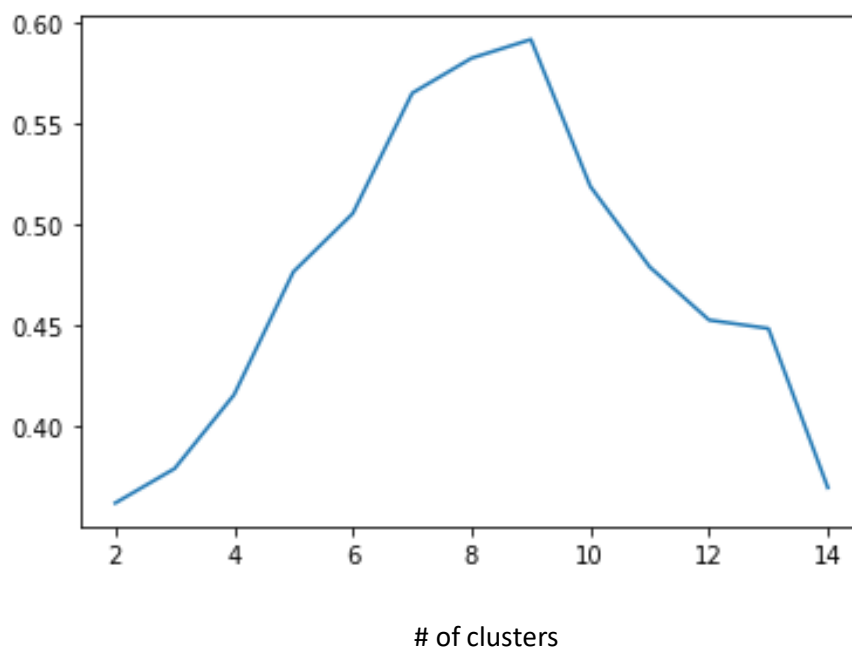
African restaurant	Airport	American Restaurant	Antique Shop	Arcade	Arepa Restaurant	...	Weight Loss Center	Whisky Bar	Wine Bar	Wings Joint	Women's Store	Yoga Studio	Community Area	Crime Score	Eatery score	Shop score
0.000000	0.0	0.093750	0.000000	0.000000	0.0	...	0.0	0.0	0.0	0.000000	0.000000	0.000000	1	0	1	1
0.000000	0.0	0.025000	0.000000	0.000000	0.0	...	0.0	0.0	0.0	0.000000	0.000000	0.000000	2	0	1	1
0.012048	0.0	0.000000	0.000000	0.012048	0.0	...	0.0	0.0	0.0	0.012048	0.000000	0.012048	3	0	1	1
0.000000	0.0	0.025316	0.000000	0.000000	0.0	...	0.0	0.0	0.0	0.000000	0.000000	0.012658	4	1	1	1
0.000000	0.0	0.033333	0.000000	0.000000	0.0	...	0.0	0.0	0.0	0.000000	0.000000	0.033333	5	1	1	1
...
0.000000	0.0	0.000000	0.000000	0.000000	0.0	...	0.0	0.0	0.0	0.000000	0.000000	0.000000	73	1	0	0
0.000000	0.0	0.000000	0.000000	0.000000	0.0	...	0.0	0.0	0.0	0.000000	0.142857	0.000000	74	1	0	0
0.000000	0.0	0.000000	0.000000	0.000000	0.0	...	0.0	0.0	0.0	0.000000	0.000000	0.000000	75	1	0	0

A bar chart is created to visualize the total scores:



3.3 Predictive Modeling

K means clustering algorithm will be used to cluster the neighborhoods into various clusters. The number of clusters to be used is determined by Silhouette coefficient for a range of 2 to 15. The number of clusters where the silhouette score reaches its highest value gives the number of clusters to be used in the model. The silhouette score diagram is shown as:



From the diagram, it is evident that silhouette score reaches high value when the # of clusters is 9. So, 9 clusters of neighborhoods are formed and the scores are analyzed per cluster.

4 Results

The cluster labels are assigned to each neighborhood for further analysis.

Total crimes	Crime Score	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	Cluster Labels	Eatery score	Shop score	Total score
3080	0	Mexican Restaurant	American Restaurant	Bakery	Asian Restaurant	Pizza Place	Bar	Theater	Farmers Market	Sandwich Place	Sushi Restaurant	4.0	1	1	2
2878	0	Indian Restaurant	Pakistani Restaurant	Grocery Store	Athletics & Sports	Donut Shop	Clothing Store	Fast Food Restaurant	Chinese Restaurant	Burmese Restaurant	Football Stadium	4.0	1	1	2
2746	0	Coffee Shop	Pizza Place	Mexican Restaurant	Chinese Restaurant	Bar	Sushi Restaurant	Diner	Thai Restaurant	Sandwich Place	Discount Store	4.0	1	1	2
1751	1	Bar	Sandwich Place	Café	Thai Restaurant	Mobile Phone Shop	Sushi Restaurant	Chinese Restaurant	Art Gallery	Pharmacy	American Restaurant	1.0	1	1	3
1055	1	Bar	Pub	Boutique	Coffee Shop	Mobile Phone Shop	Diner	Bakery	Chinese Restaurant	Music Venue	Dance Studio	1.0	1	1	3

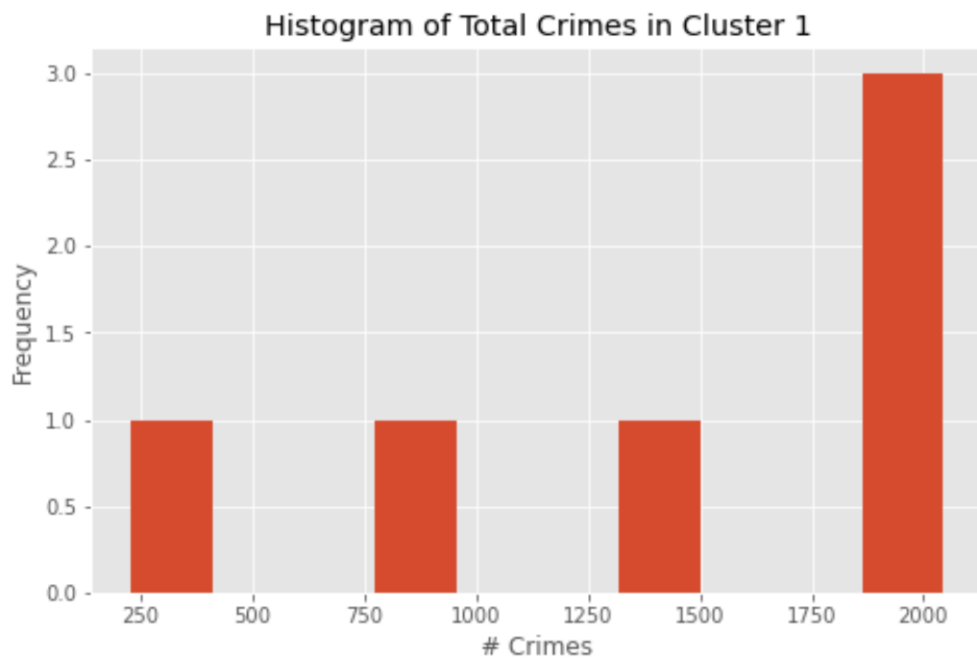
Cluster 1

There are 6 neighborhoods in cluster 1,

	Community Name	Total crimes	Crime Score	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	Cluster Labels
8	Edison Park	229	1	Bar	Italian Restaurant	Mexican Restaurant	Theater	Track	Grocery Store	French Restaurant	Salon / Barbershop	Pub	Spa	0.0
13	Albany Park	1901	1	Mexican Restaurant	Korean Restaurant	Grocery Store	Park	Asian Restaurant	Latin American Restaurant	Sushi Restaurant	Liquor Store	Taco Place	Karaoke Bar	0.0
15	Irving Park	2044	1	Bar	Dance Studio	American Restaurant	Asian Restaurant	Moroccan Restaurant	Latin American Restaurant	Furniture / Home Store	Supermarket	Breakfast Spot	Bus Station	0.0
32	Near South Side	1477	1	Football Stadium	Athletics & Sports	Historic Site	Park	Bakery	New American Restaurant	Bar	Salon / Barbershop	Cocktail Bar	Sushi Restaurant	0.0
33	Armour Square	797	1	Chinese Restaurant	Pizza Place	Mexican Restaurant	Asian Restaurant	Szechuan Restaurant	Business Service	Park	Storage Facility	Mobile Phone Shop	Flea Market	0.0
34	Douglas	2026	1	Fast Food Restaurant	Fried Chicken Joint	Wings Joint	Cosmetics Shop	Sandwich Place	Mobile Phone Shop	Residential Building (Apartment)	Liquor Store	Eye Doctor	Bank	0.0

	Community Name	Total crimes	Eatery score	Shop score	Total score
8	Edison Park	229	1	0	2
13	Albany Park	1901	1	0	2
15	Irving Park	2044	1	0	2
32	Near South Side	1477	1	0	2
33	Armour Square	797	1	0	2
34	Douglas	2026	1	0	2

Histogram of the total crimes for these neighborhood



Histogram of the total scores for cluster 1



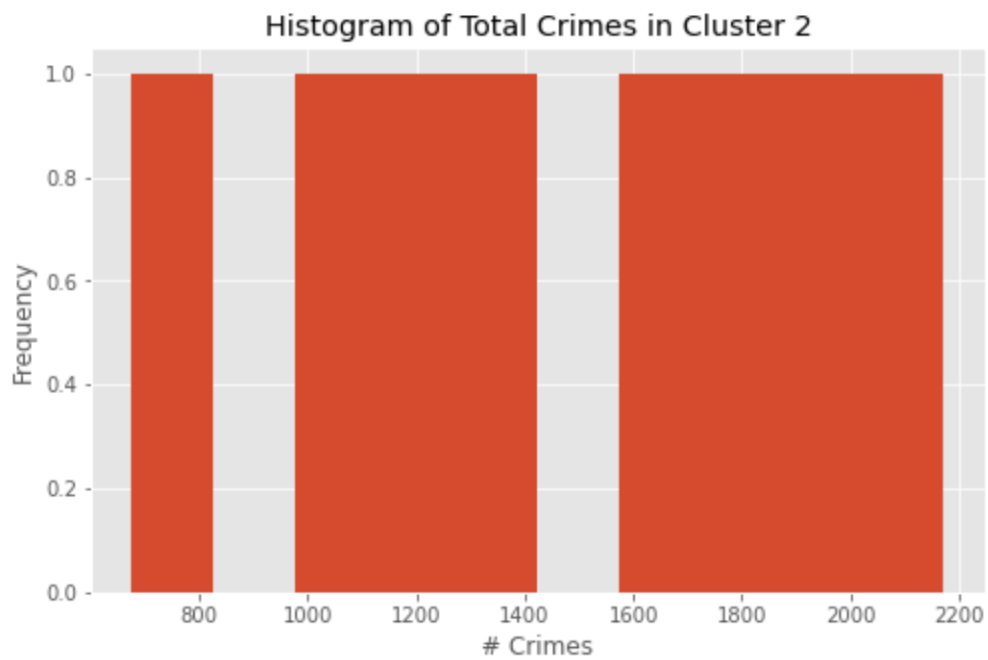
Cluster 2

There are 8 neighborhoods in cluster 2,

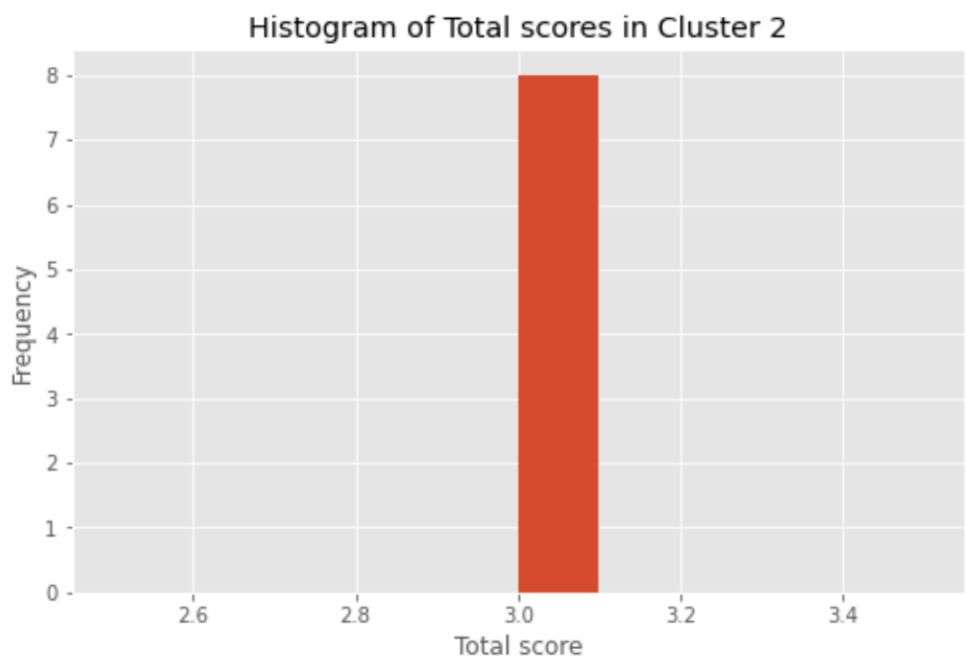
	Community Name	Total crimes	Crime Score	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	Cluster Labels
3	Lincoln Square	1751	1	Bar	Sandwich Place	Café	Thai Restaurant	Mobile Phone Shop	Sushi Restaurant	Chinese Restaurant	Art Gallery	Pharmacy	American Restaurant	1.0
4	North Center	1055	1	Bar	Pub	Boutique	Coffee Shop	Mobile Phone Shop	Diner	Bakery	Chinese Restaurant	Music Venue	Dance Studio	1.0
20	Avondale	1689	1	Park	Hot Dog Joint	Bar	Chinese Restaurant	Mexican Restaurant	Breakfast Spot	ATM	Food Truck	Korean Restaurant	Supermarket	1.0
30	Lower West Side	1906	1	Mexican Restaurant	Pizza Place	Bakery	Thrift / Vintage Store	Latin American Restaurant	Bar	Plaza	Taco Place	Breakfast Spot	Massage Studio	1.0
40	Hyde Park	1376	1	Sandwich Place	Mobile Phone Shop	Thai Restaurant	Café	Pizza Place	American Restaurant	Cosmetics Shop	Ice Cream Shop	Bookstore	Coffee Shop	1.0
59	Bridgeport	1180	1	Chinese Restaurant	Bar	Pizza Place	Mexican Restaurant	American Restaurant	Grocery Store	Wings Joint	Korean Restaurant	Bus Station	Coffee Shop	1.0
61	West Elsdon	676	1	Mexican Restaurant	Chinese Restaurant	Pizza Place	Seafood Restaurant	Supplement Shop	Supermarket	Donut Shop	Grocery Store	Fried Chicken Joint	Gas Station	1.0

	Community Name	Total crimes	Eatery score	Shop score	Total score
3	Lincoln Square	1751	1	1	3
4	North Center	1055	1	1	3
20	Avondale	1689	1	1	3
30	Lower West Side	1906	1	1	3
40	Hyde Park	1376	1	1	3
59	Bridgeport	1180	1	1	3
61	West Elsdon	676	1	1	3
75	Edgewater	2172	1	1	3

The total crimes for these neighborhood is captured in a histogram.



Histogram of the total scores for cluster 2



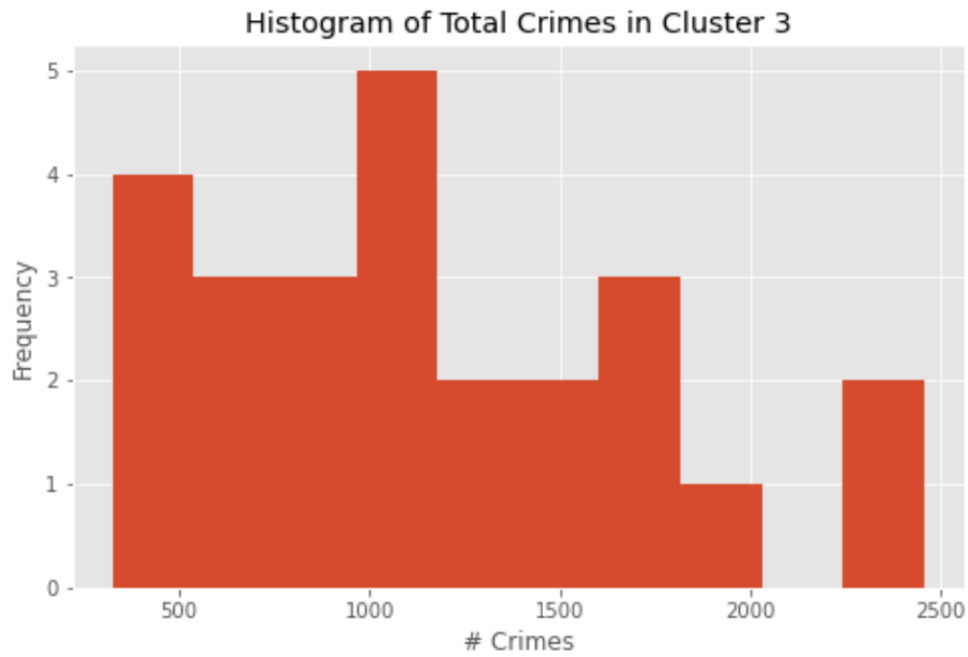
Cluster 3

There are 25 neighborhoods in cluster 3,

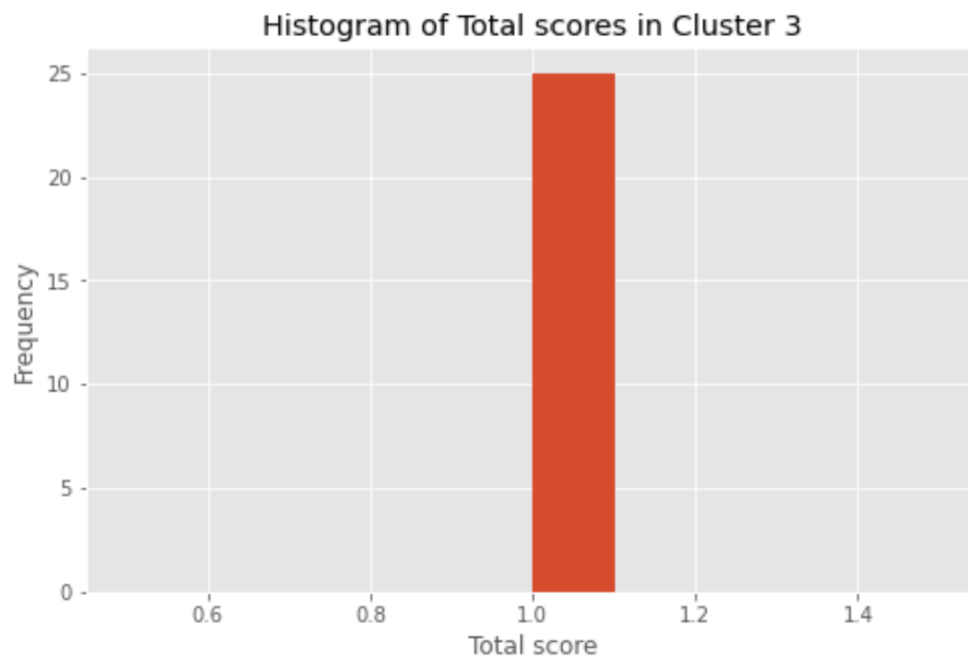
Community Name	Total crimes	Crime Score	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	Cluster Labels	Eatery score	Si sc
Norwood Park	893	1	Train Station	Sandwich Place	Hobby Shop	Spa	Gastropub	American Restaurant	Automotive Shop	Italian Restaurant	Thai Restaurant	Diner	2.0	0	
Forest Glen	496	1	Other Great Outdoors	Playground	Soccer Field	Golf Course	Yoga Studio	Fast Food Restaurant	Eye Doctor	Falafel Restaurant	Farm	Farmers Market	2.0	0	
North Park	720	1	Korean Restaurant	Coffee Shop	Convenience Store	College Quad	Food Truck	Pizza Place	Bus Station	Sandwich Place	Taco Place	Bakery	2.0	0	
Portage Park	2430	1	Mexican Restaurant	Pizza Place	Playground	Sports Bar	Dog Run	Rental Car Location	Coffee Shop	Park	Track	Farmers Market	2.0	0	
Dunning	1229	1	Thai Restaurant	Polish Restaurant	Pet Store	Tattoo Parlor	Fast Food Restaurant	Asian Restaurant	Japanese Restaurant	Gym Pool	Deli / Bodega	Spa	2.0	0	

	Community Name	Total crimes	Eatery score	Shop score	Total score
9	Norwood Park	893	0	0	1
11	Forest Glen	496	0	0	1
12	North Park	720	0	0	1
14	Portage Park	2430	0	0	1
16	Dunning	1229	0	0	1
19	Hermosa	1107	0	0	1
35	Oakland	575	0	0	1
36	Fuller Park	646	0	0	1
38	Kenwood	1123	0	0	1
39	Washington Park	1834	0	0	1
44	Avalon Park	1129	0	0	1
46	Burnside	327	0	0	1
47	Calumet Heights	1191	0	0	1
49	Pullman	871	0	0	1
51	East Side	881	0	0	1
54	Hegewisch	468	0	0	1
55	Garfield Ridge	1436	0	0	1
57	Brighton Park	1667	0	0	1
62	Gage Park	1608	0	0	1
64	West Lawn	1145	0	0	1
69	Ashburn	1753	0	0	1
71	Washington Heights	2455	0	0	1
72	Mount Greenwood	456	0	0	1
73	Morgan Park	1578	0	0	1
74	O'Hare	1053	0	0	1

Histogram showing the total crimes for these neighborhoods.



Histogram of the total scores for cluster 3



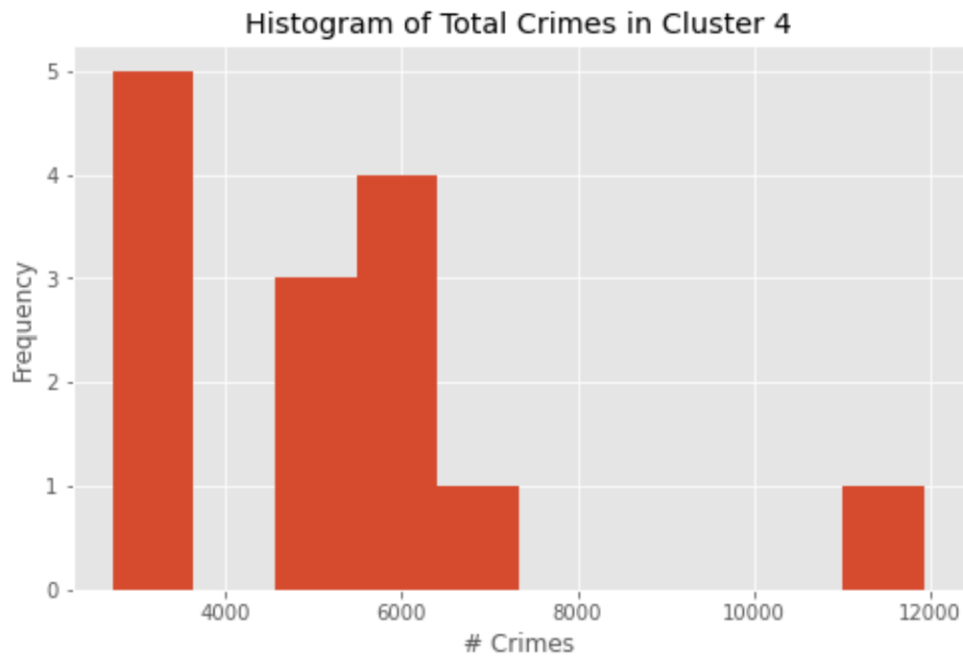
Cluster 4

There are 14 neighborhoods in cluster 4,

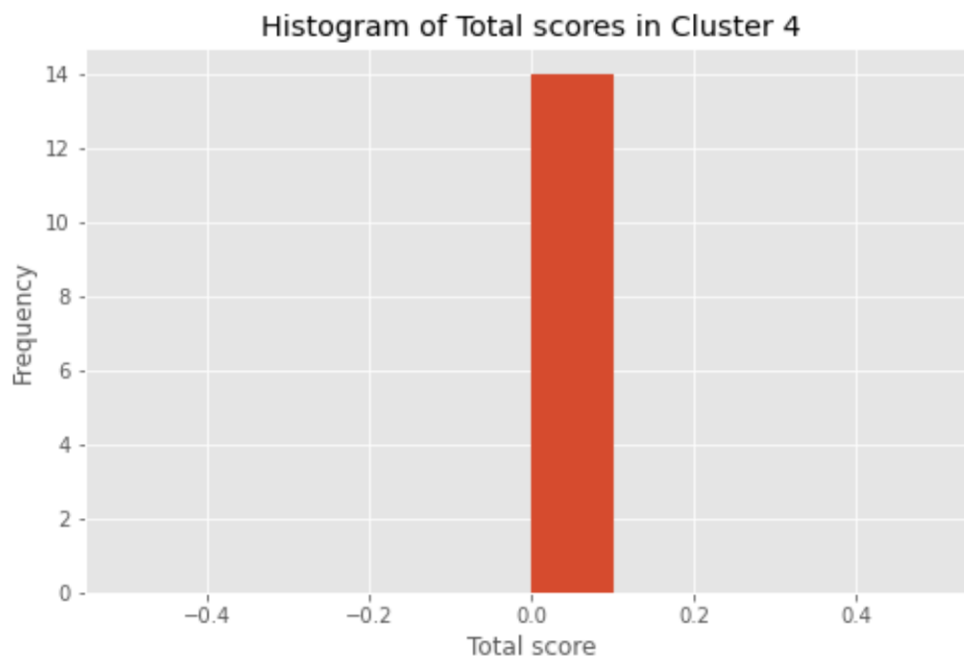
Community Name	Total crimes	Crime Score	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	Cluster Labels
Belmont Cragin	3502	0	Mexican Restaurant	Convenience Store	Donut Shop	Ice Cream Shop	American Restaurant	Grocery Store	Athletics & Sports	Sandwich Place	Supermarket	Field	3.0
Humboldt Park	6278	0	Park	Liquor Store	Taco Place	Donut Shop	Grocery Store	Video Store	Art Gallery	Event Service	Yoga Studio	Farm	3.0
Austin	11928	0	Discount Store	Cosmetics Shop	Breakfast Spot	Pizza Place	Salon / Barbershop	Café	Field	Farm	Farmers Market	Fast Food Restaurant	3.0
East Garfield Park	3534	0	Discount Store	Diner	Bus Station	Bus Line	Supermarket	Liquor Store	Pharmacy	Southern / Soul Food Restaurant	Floating Market	Flea Market	3.0
North Lawndale	6675	0	Fast Food Restaurant	BBQ Joint	Convenience Store	Yoga Studio	Food Court	Food & Drink Shop	Food	Flower Shop	Floating Market	Flea Market	3.0
Woodlawn	2720	0	History Museum	Coffee Shop	Bookstore	Indie Movie Theater	Pub	Café	Market	College Cafeteria	Park	Sandwich Place	3.0
Chatham	5139	0	Park	Boutique	Lounge	Discount Store	Donut Shop	Fast Food Restaurant	Food	Fried Chicken Joint	Ice Cream Shop	Hardware Store	3.0
South Chicago	3168	0	Park	Mexican Restaurant	Pizza Place	American Restaurant	Filipino Restaurant	Falafel Restaurant	Farm	Farmers Market	Fast Food Restaurant	Field	3.0

	Community Name	Total crimes	Eatery score	Shop score
18	Belmont Cragin	3502	0	0
22	Humboldt Park	6278	0	0
24	Austin	11928	0	0
26	East Garfield Park	3534	0	0
28	North Lawndale	6675	0	0
41	Woodlawn	2720	0	0
43	Chatham	5139	0	0
45	South Chicago	3168	0	0
48	Roseland	5600	0	0
52	West Pullman	3628	0	0
66	West Englewood	5369	0	0
67	Englewood	4874	0	0
68	Greater Grand Crossing	5579	0	0
70	Auburn Gresham	5852	0	0

Histogram showing the total crimes for these neighborhoods.



Histogram of the total scores for cluster 4



Cluster 5

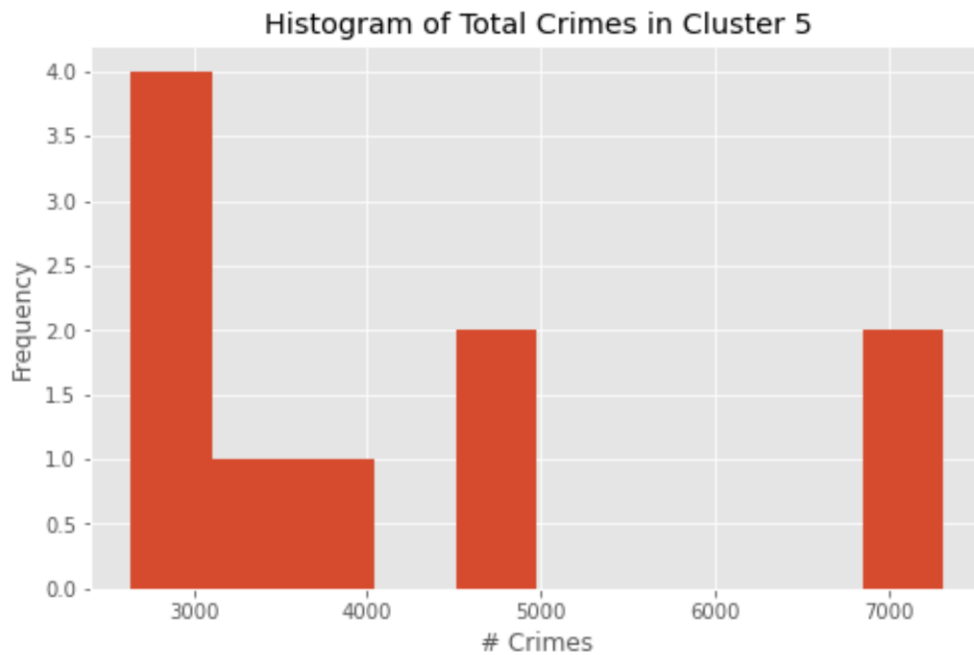
There are 10 neighborhoods in cluster 5,

Community Name	Total crimes	Crime Score	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	Cluster Labels
Rogers Park	3080	0	Mexican Restaurant	American Restaurant	Bakery	Asian Restaurant	Pizza Place	Bar	Theater	Farmers Market	Sandwich Place	Sushi Restaurant	4.0
West Ridge	2878	0	Indian Restaurant	Pakistani Restaurant	Grocery Store	Athletics & Sports	Donut Shop	Clothing Store	Fast Food Restaurant	Chinese Restaurant	Burmese Restaurant	Football Stadium	4.0
Uptown	2746	0	Coffee Shop	Pizza Place	Mexican Restaurant	Chinese Restaurant	Bar	Sushi Restaurant	Diner	Thai Restaurant	Sandwich Place	Discount Store	4.0
Lake View	4007	0	Sandwich Place	Pizza Place	Coffee Shop	Bagel Shop	Bakery	Bar	Performing Arts Venue	Spa	Café	Pub	4.0
Lincoln Park	2642	0	Bar	Coffee Shop	Pizza Place	Sandwich Place	Yoga Studio	Mexican Restaurant	Theater	Taco Place	Ice Cream Shop	Hot Dog Joint	4.0
Near North Side	6859	0	Hotel	Steakhouse	Bar	Italian Restaurant	Clothing Store	Boutique	Coffee Shop	Gym / Fitness Center	American Restaurant	Lounge	4.0
Logan Square	3480	0	Cocktail Bar	Bar	Mexican Restaurant	Coffee Shop	Pizza Place	Brewery	Café	Yoga Studio	Bike Shop	Bowling Alley	4.0
West Town	4957	0	Bar	Café	Salon / Barbershop	Pizza Place	Sushi Restaurant	Bakery	Argentinian Restaurant	Pet Store	Cocktail Bar	Italian Restaurant	4.0
Loop	4843	0	Coffee Shop	American Restaurant	Hotel	Theater	Bakery	Museum	Snack Place	Department Store	Plaza	Bar	4.0

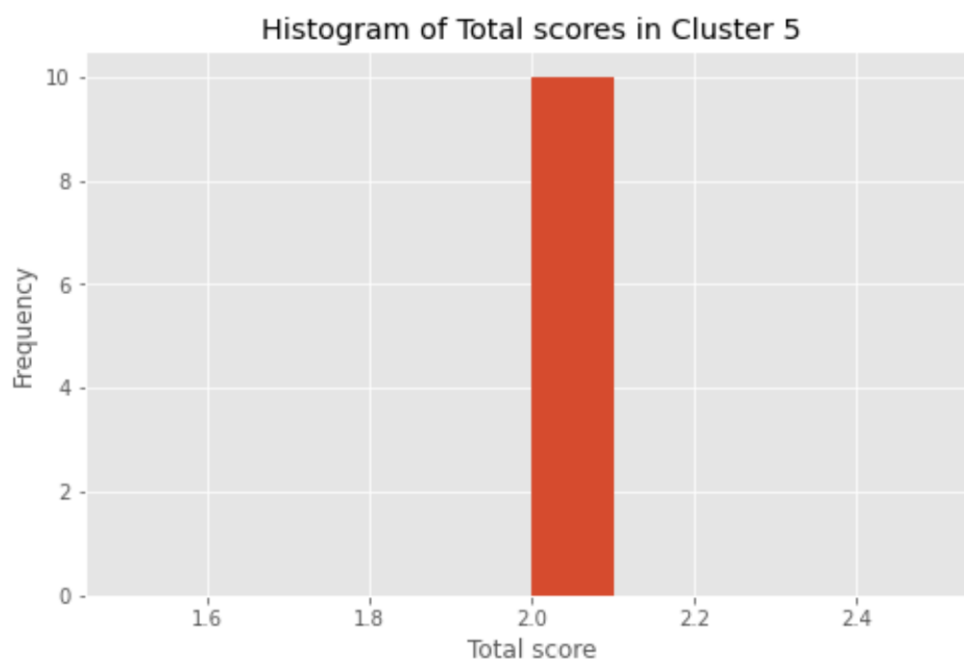
End

	Community Name	Total crimes	Eatery score	Shop score	Total score
0	Rogers Park	3080	1	1	2
1	West Ridge	2878	1	1	2
2	Uptown	2746	1	1	2
5	Lake View	4007	1	1	2
6	Lincoln Park	2642	1	1	2
7	Near North Side	6859	1	1	2
21	Logan Square	3480	1	1	2
23	West Town	4957	1	1	2
31	Loop	4843	1	1	2
42	South Shore	7320	1	1	2

Histogram showing the total crimes for these neighborhoods.



Histogram of the total scores for cluster 5



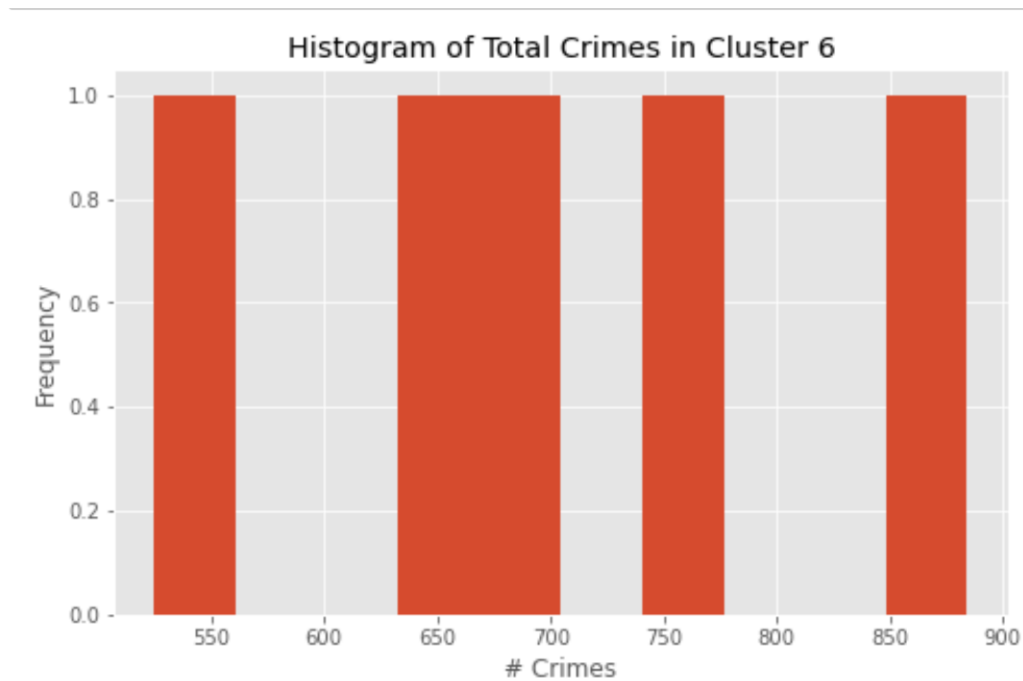
Cluster 6

There are 5 neighborhoods in cluster 6,

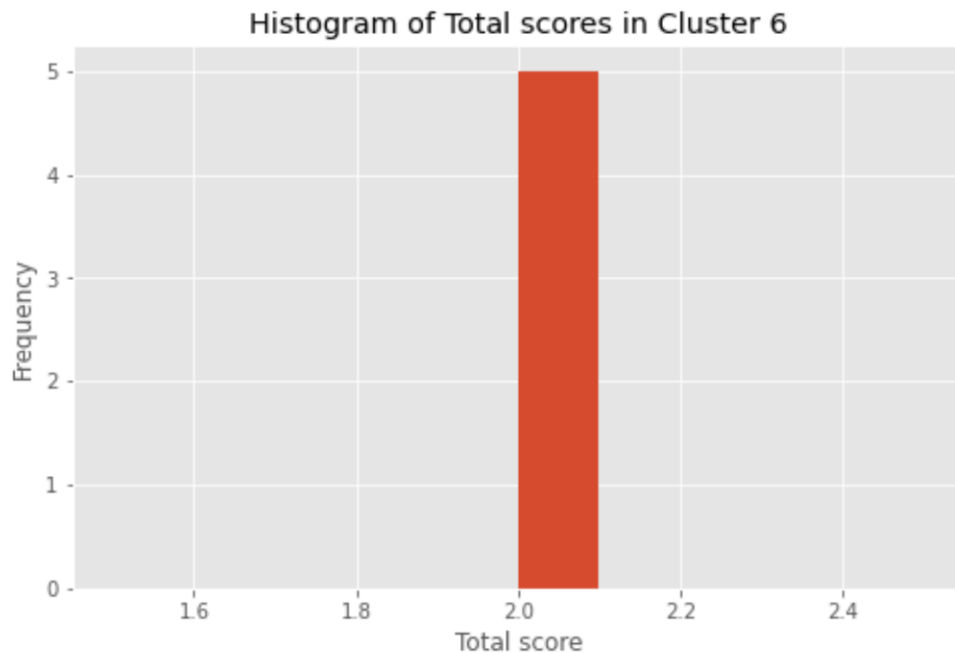
Community Name	Total crimes	Crime Score	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	Cluster Labels
Jefferson Park	884	1	Discount Store	Chinese Restaurant	Mobile Phone Shop	Bar	Market	Automotive Shop	Greek Restaurant	Sandwich Place	Dive Bar	Bowling Alley	5.0
Montclare	525	1	Mexican Restaurant	Furniture / Home Store	Bus Station	Pizza Place	Breakfast Spot	Café	Other Repair Shop	Electronics Store	Automotive Shop	Grocery Store	5.0
Archer Heights	656	1	Mexican Restaurant	Discount Store	Sandwich Place	Bank	Shopping Mall	Shoe Store	Seafood Restaurant	Mobile Phone Shop	Supplement Shop	Rental Car Location	5.0
McKinley Park	697	1	Mexican Restaurant	Video Store	Thrift / Vintage Store	Diner	Grocery Store	Baseball Field	Supermarket	Soccer Field	Donut Shop	Bus Station	5.0
Clearing	772	1	Construction & Landscaping	Pizza Place	Video Store	Playground	Donut Shop	Discount Store	Record Shop	Grocery Store	Fast Food Restaurant	Chinese Restaurant	5.0

	Community Name	Total crimes	Eatery score	Shop score	Total score
10	Jefferson Park	884	0	1	2
17	Montclare	525	0	1	2
56	Archer Heights	656	0	1	2
58	McKinley Park	697	0	1	2
63	Clearing	772	0	1	2

Histogram showing the total crimes for these neighborhoods.



Histogram of the total scores for cluster 6



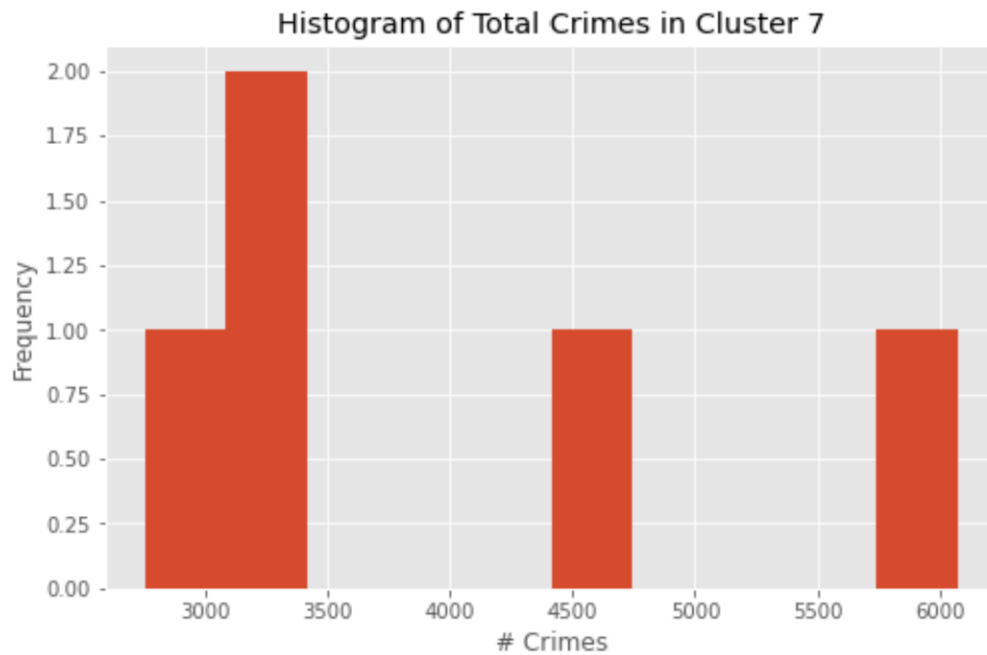
Cluster 7

There are 5 neighborhoods in cluster 7,

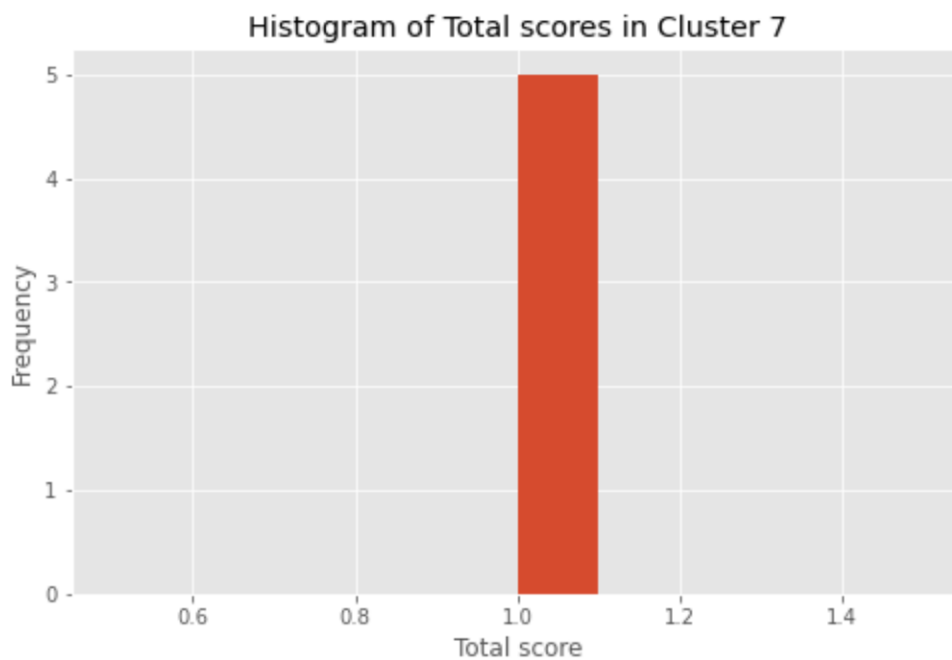
Community Name	Total crimes	Crime Score	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	Cluster Labels	Eatery score
West Garfield Park	4681	0	Fast Food Restaurant	Shoe Store	Fried Chicken Joint	Clothing Store	Liquor Store	Caribbean Restaurant	Mobile Phone Shop	Kids Store	Supermarket	Sandwich Place	6.0	0
Near West Side	6072	0	Park	Café	Dance Studio	Coffee Shop	Bar	Yoga Studio	College Stadium	Bus Station	Sports Bar	Stadium	6.0	0
South Lawndale	3354	0	Mexican Restaurant	Mobile Phone Shop	Park	Food & Drink Shop	Supermarket	Liquor Store	Optical Shop	Bakery	Restaurant	Grocery Store	6.0	0
Grand Boulevard	2757	0	Liquor Store	Art Gallery	Cosmetics Shop	BBQ Joint	Burger Joint	Shoe Store	Breakfast Spot	Caribbean Restaurant	Sandwich Place	Lounge	6.0	0
New City	3373	0	Pizza Place	Grocery Store	Bank	Brewery	Mexican Restaurant	Department Store	Pharmacy	Shoe Store	Farm	Chinese Restaurant	6.0	0

	Community Name	Total crimes	Eatery score	Shop score	Total score
25	West Garfield Park	4681	0	1	1
27	Near West Side	6072	0	1	1
29	South Lawndale	3354	0	1	1
37	Grand Boulevard	2757	0	1	1
60	New City	3373	0	1	1

Histogram showing the total crimes for these neighborhoods.



Histogram of the total scores for cluster 7



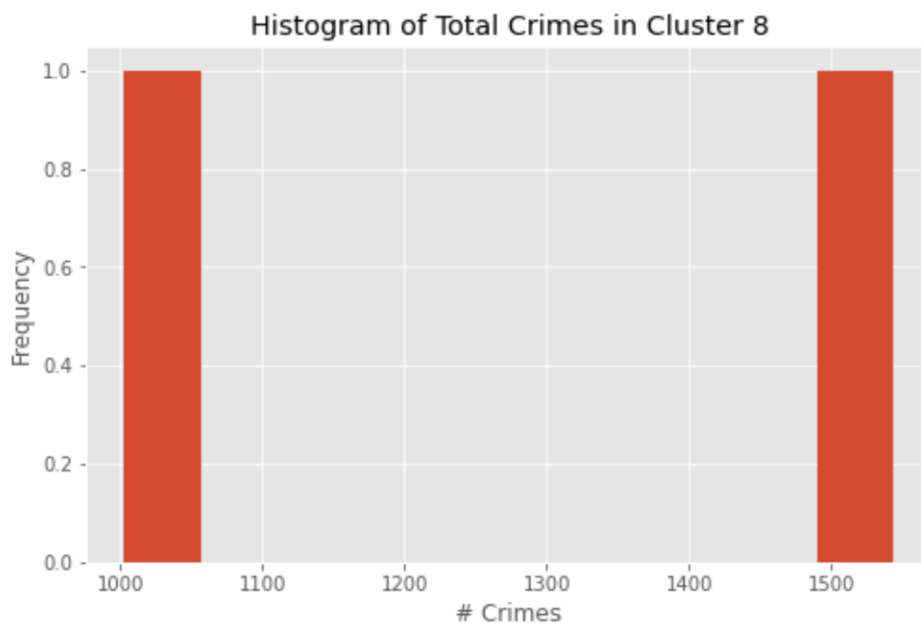
Cluster 8

There are 2 neighborhoods in cluster 8,

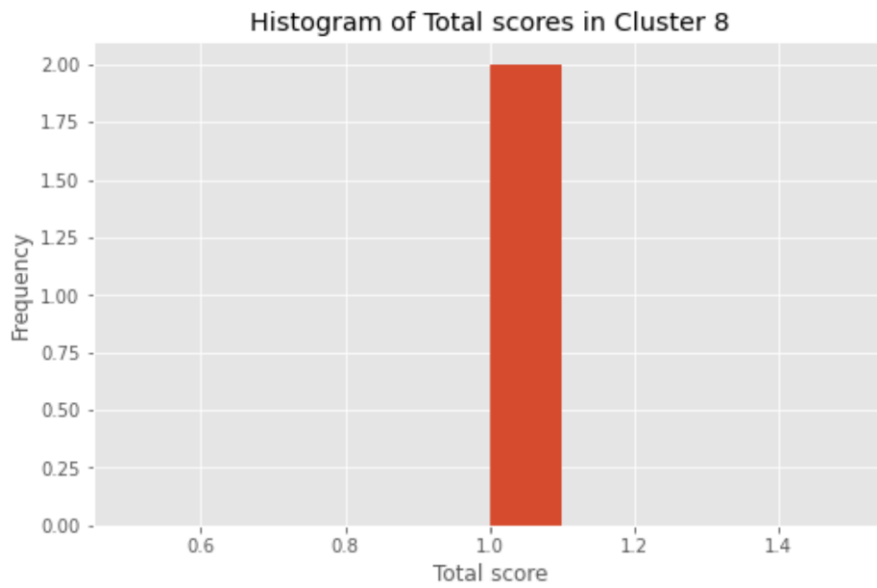
	Community Name	Total crimes	Crime Score	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	Cluster Labels	Eatery score	S score
50	South Deering	1544	1	Park	Yoga Studio	Event Service	Food Court	Food & Drink Shop	Food	Flower Shop	Floating Market	Flea Market	Filipino Restaurant	7.0	0	
53	Riverdale	1003	1	Park	Yoga Studio	Event Service	Food Court	Food & Drink Shop	Food	Flower Shop	Floating Market	Flea Market	Filipino Restaurant	7.0	0	

	Community Name	Total crimes	Eatery score	Shop score	Total score
50	South Deering	1544	0	0	1
53	Riverdale	1003	0	0	1

Histogram showing the total crimes for these neighborhoods.



Histogram of the total scores for cluster 8



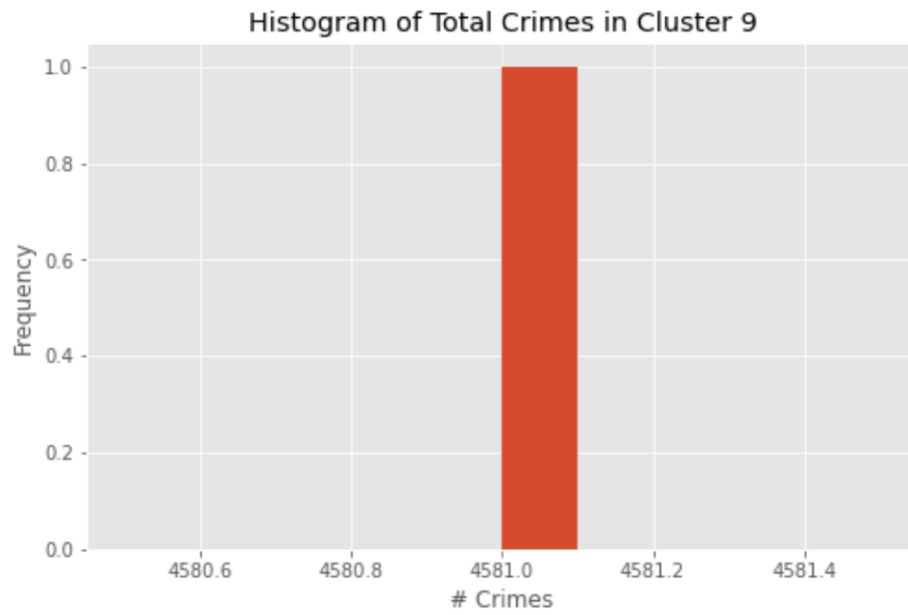
Cluster 9

There is 1 neighborhood in cluster 9,

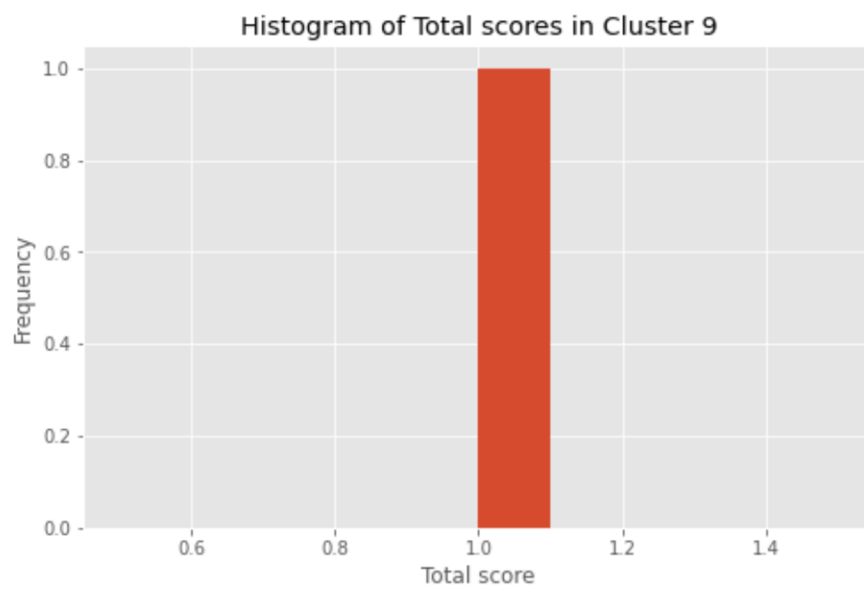
Community Name	Total crimes	Crime Score	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	Cluster Labels	Eatery score
Chicago Lawn	4581	0	Mexican Restaurant	Pizza Place	Cosmetics Shop	Fast Food Restaurant	Intersection	American Restaurant	Flea Market	Farm	Farmers Market	Field	8.0	1

Community Name	Total crimes	Eatery score	Shop score	Total score
65 Chicago Lawn	4581	1	0	1

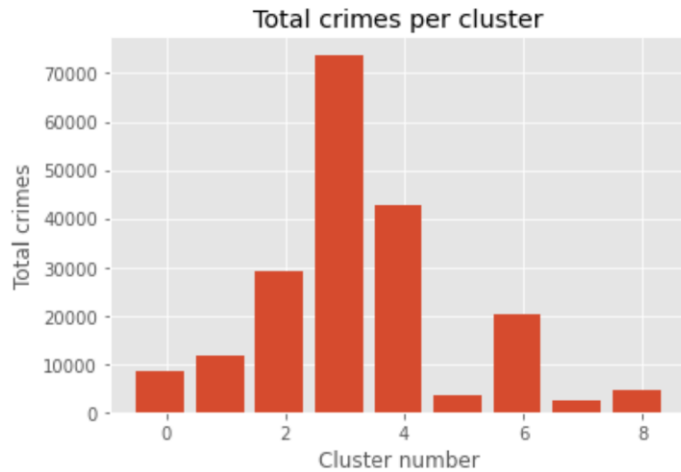
Histogram showing the total crimes for these neighborhoods.



Histogram of the total scores for cluster 9



Total crimes per cluster



5 Discussion

The average number of crimes in the entire neighborhoods is approximately 2590. In the first cluster, almost all the neighborhoods show the crimes less than the average and the number of eatery places are more than the average. But, the number of shopping places is less than the average # of shopping centers around. Edison park and Armour square shows less crime incidents.

In the second cluster, the crimes are less than the average of the total crimes and the shopping and eatery places are more than the average of the total number of centers around. With all the shopping and restaurants around, West Elsdon has total number of crimes less than 1000, which makes it a good neighborhood comparatively. In cluster 3, the total score is simply 1 with the crimes less than the average, but less shopping places and restaurants. This makes it less desirable cluster of neighborhoods. In cluster 4, the total score is 0, with crimes more than the average and shopping places and restaurants less than

the average. In cluster 5, the crime rate is high compared to other clusters. Even though there are enough eatery and shopping places, it is not a preferrable cluster of neighborhoods to move in.

In cluster 6, the crime rate is comparatively low compared to other clusters, but not enough restaurants around the neighborhoods. In cluster 7, the crime rate is higher than the average crime incidents, which makes it less preferrable cluster of neighborhoods. In cluster 8, there are not enough restaurants and shopping places around, with the total score of 1, which also makes it less preferable cluster 9, the crime rate is high and not enough shopping places around the neighborhood.

Out of the 9 clusters, the second cluster with the total score of 3 makes it the best cluster to look into. Out of that, West Elsdon has a crime rate less than 1000 and surrounded by restaurants and eatery places more than the average available in all of the neighborhoods collectively. Then clusters 1,5,6 has a total score of 2, with not enough eatery and shopping places around and cluster 5, the crime incidents are more than the average. Cluster 6 has the least number of crime incidents comparatively. Clusters 3,7,8,9 has a score of 1, which makes it the least choice of neighborhoods, with less eatery / shopping places and more crime rates. Cluster 4 is the least preferrable neighborhood with a total score of 0, as it has a high crime rate and not enough eatery and shopping centers around.

6 Conclusion

From this project, it is found that, out of all 77 neighborhoods, West Elsdon community from cluster 2 looks the best with less crime incidents and good number of restaurants and shopping centers around. Cluster 6 has the least number of crimes comparatively. This project took into account the up-to-date crime incidents, shopping places and restaurants around the neighborhoods. The stakeholders can take up the suggestion based on these factors. However, there are also other factors involved in selecting the neighborhood like schools, parks, museums which depend on the choice of stakeholders, and is beyond the scope of this project and thereby not considered. This project can be further developed taking into account all the possible scenarios in moving to a neighborhood.