FINDING THE BEST NEIGHBORHOOD IN CHICAGO

BY

DEVI SELVAKUMAR

PURPOSE

- With the crime rate high in some Chicago neighborhoods, it is hard to find which would be a suitable neighborhood to move in.
- The objective of the project is to find the best neighborhood in Chicago, with less crime incidents and enough shopping places and eatery places around.
- The stakeholders are the people who are in search of good Chicago neighborhoods.

DATA

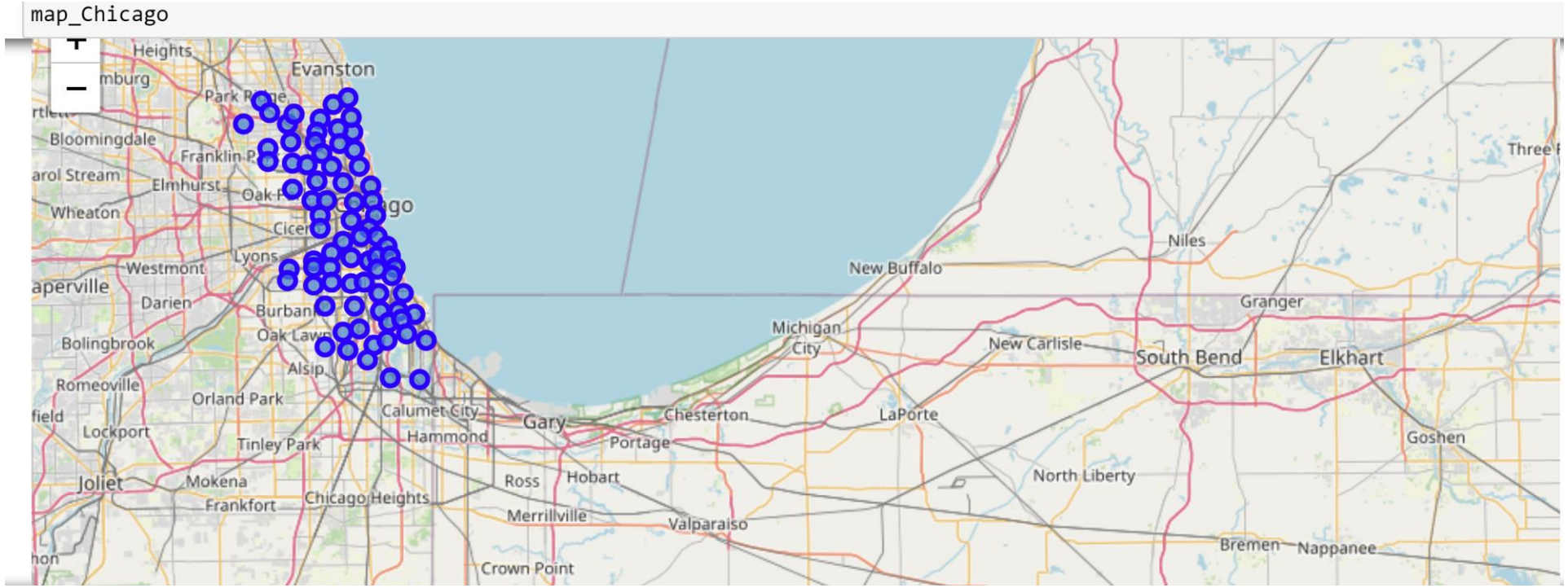
- Chicago - Current crime data set from the City of Chicago data portal https://data.cityofchicago.org/api/views/qzdf-xmn8/rows.csv It contains the crime records from Jan.1st 2020 to Dec.5th 2020.
- Chicago Community data set https://raw.githubusercontent.com/jkgiesler/parse-chicago-neighborhoods/master/community_to_gps.txt There are 77 neighborhoods in Chicago

FOURSQUARE API

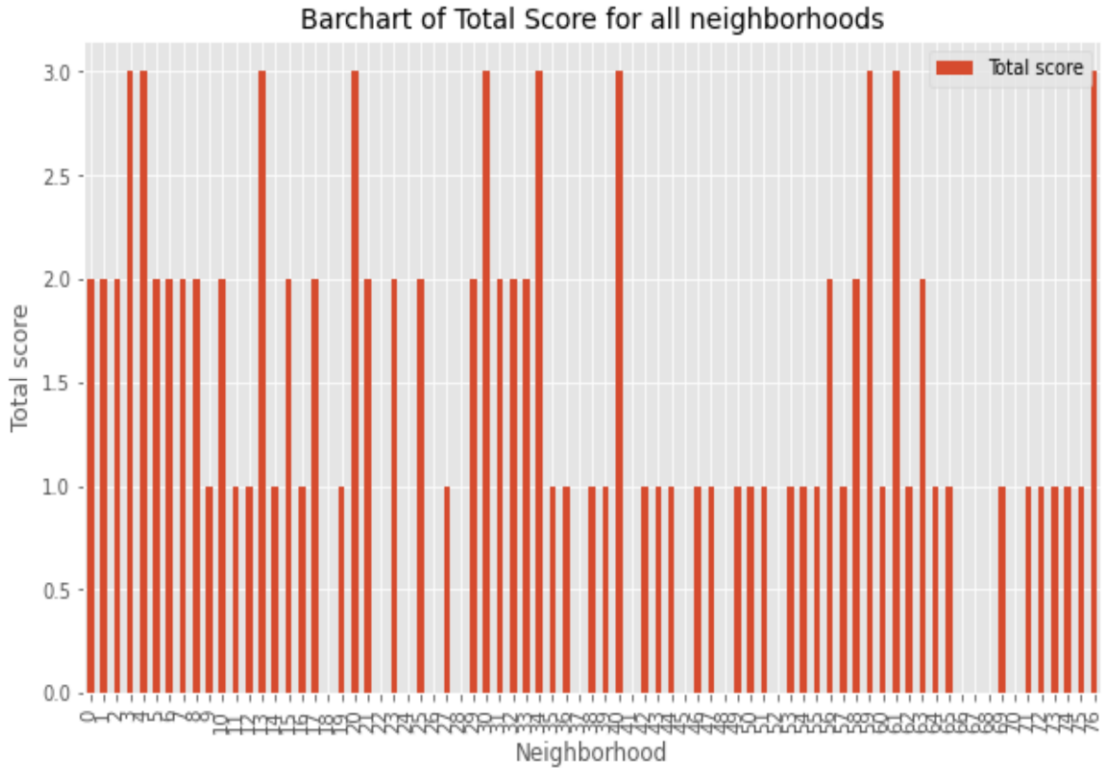
- Using Foursquare API, 10 most common venues for each neighborhood is found out:

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

CHICAGO NEIGHBORHOOD MAP



ASSIGNING SCORES



For each neighborhood, the following scores are assigned:
Crime score = 1, if the # of crimes <= average of crimes in all neighborhoods.= 0 if the # of crimes > average of crimes in all neighborhoods.

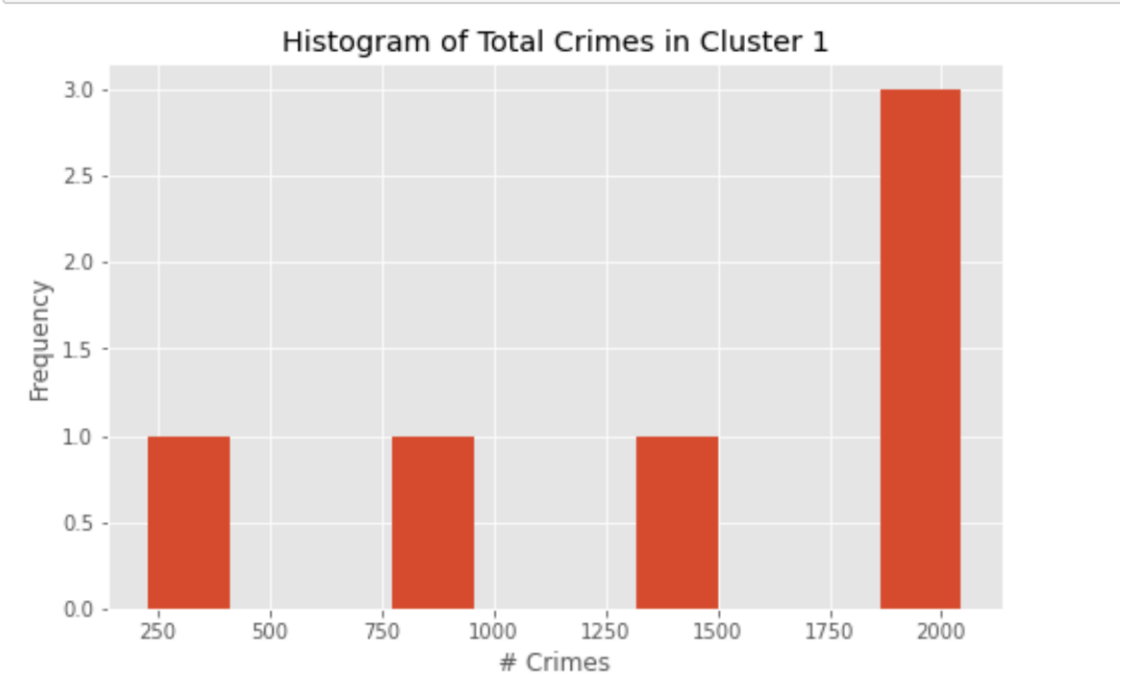
Shopping score = 1, if the # of shopping places >= average # shopping places. =0, 1, if the # of shopping places < average # shopping places.

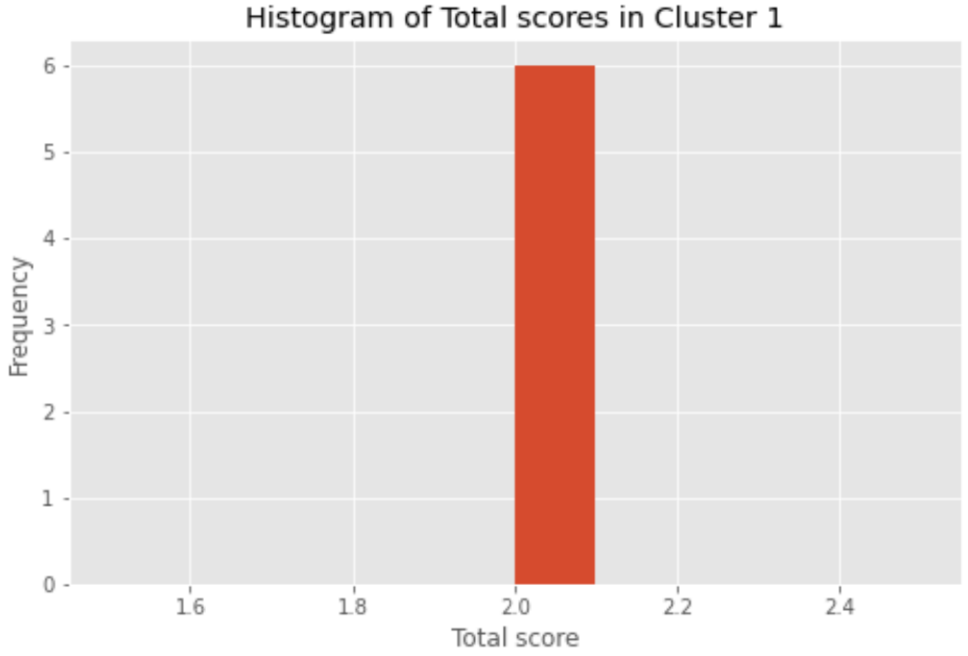
Eatery score = 1, if the # of restaurant/fast food places >= average # eatery places. =0, 1, if the # of restaurant/fast food places < average # eatery places.

Average of crimes in all neighborhoods -> approx.2500

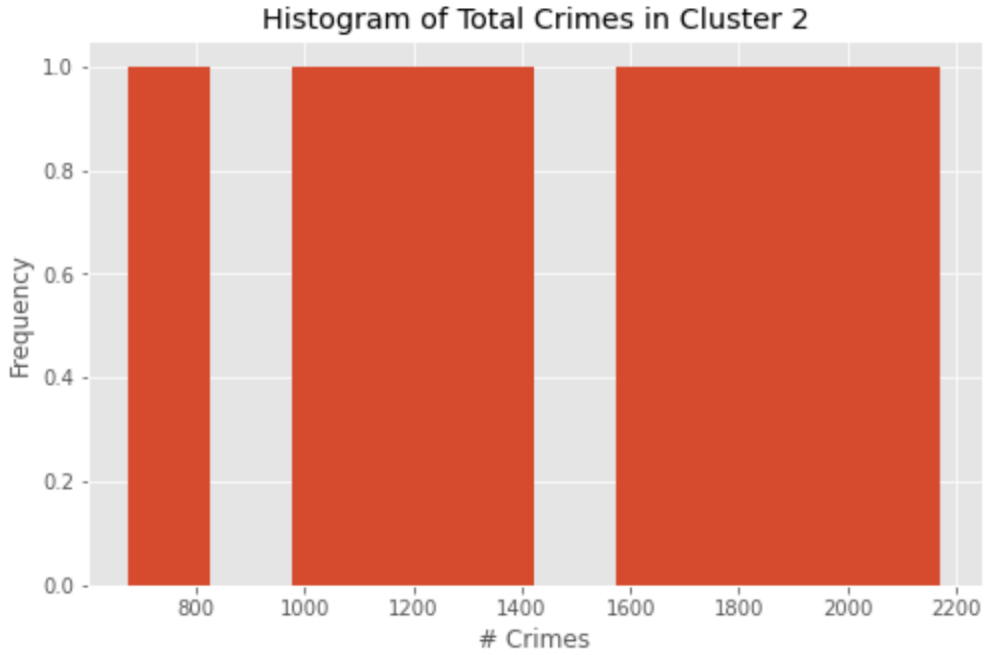
PREDICTIVE MODELING: K MEANS CLUSTERING- CLUSTER 1

- Best k= 9 clusters is determined using Silhouette score. Each cluster is analyzed for total crimes and total scores. Crime rate < average, Not enough restaurants with Total score = 2





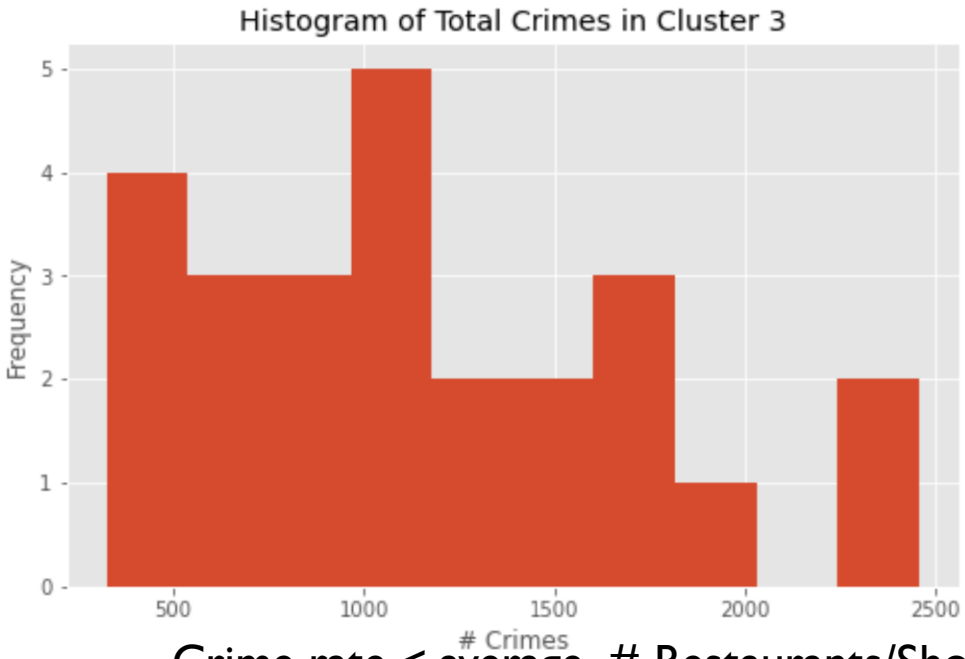
K MEANS CLUSTERING- CLUSTER 2



Crime rate < average, # Restaurants/Shopping places > average with Total score = 3 making it the best option

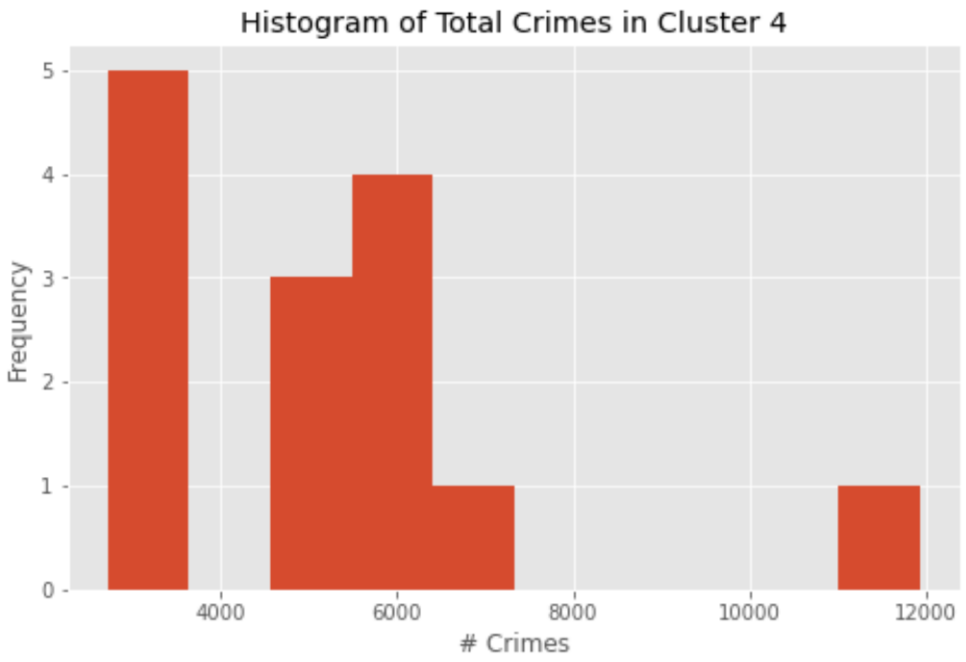
K MEANS CLUSTERING- CLUSTER 3

- Crime rate < average, # Restaurants/Shopping places < average with Total score = 1



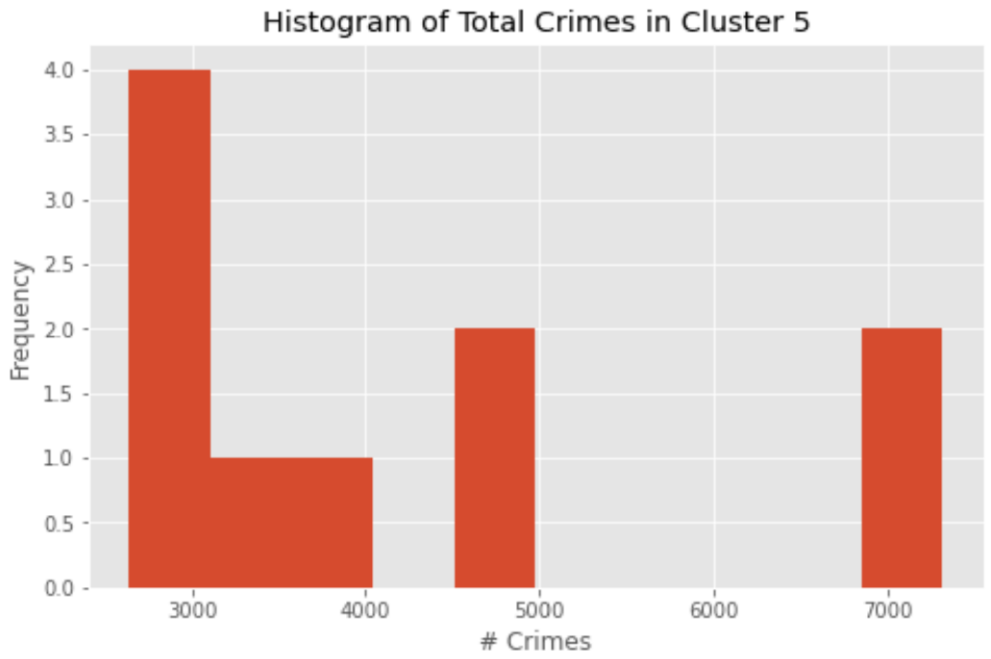
Crime rate < average, # Restaurants/Shopping places < average with Total score = 1

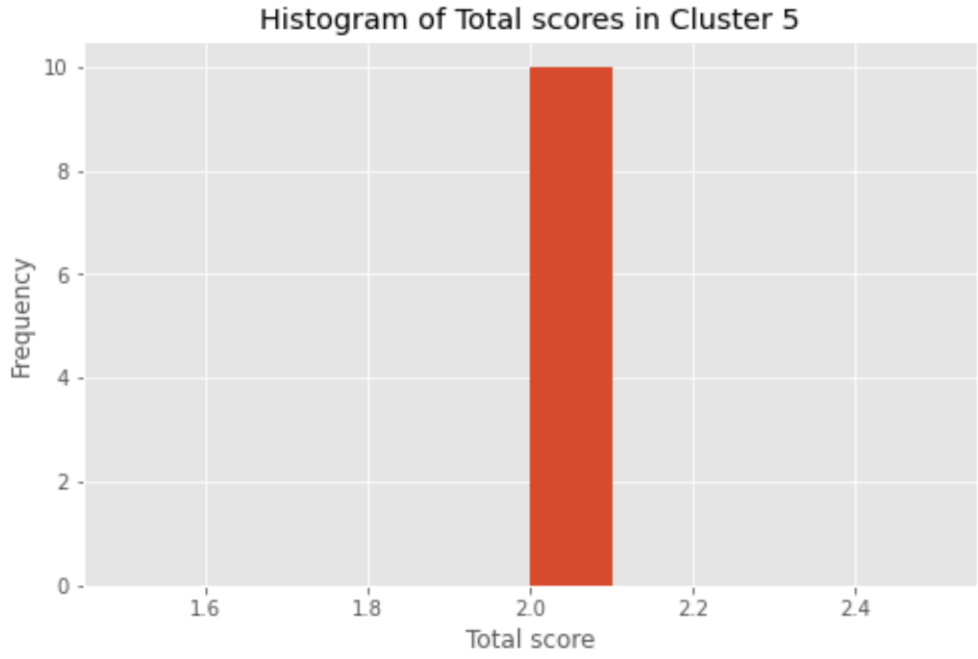
K MEANS CLUSTERING- CLUSTER 4



Crime rate > average, # Restaurants/Shopping places < average with Total score = 0 making it the least option

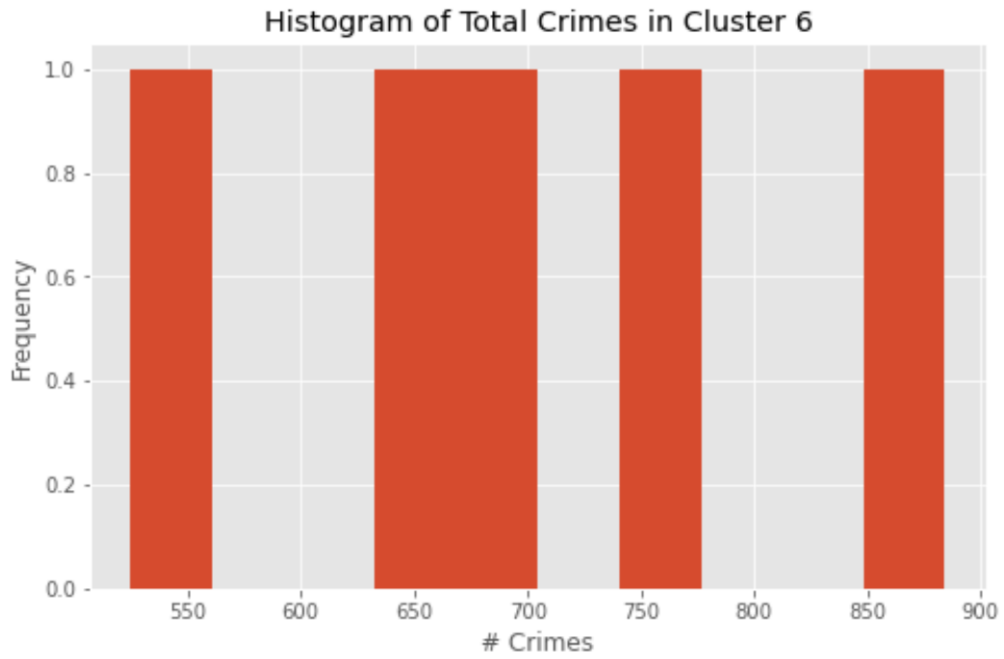
K MEANS CLUSTERING- CLUSTER 5

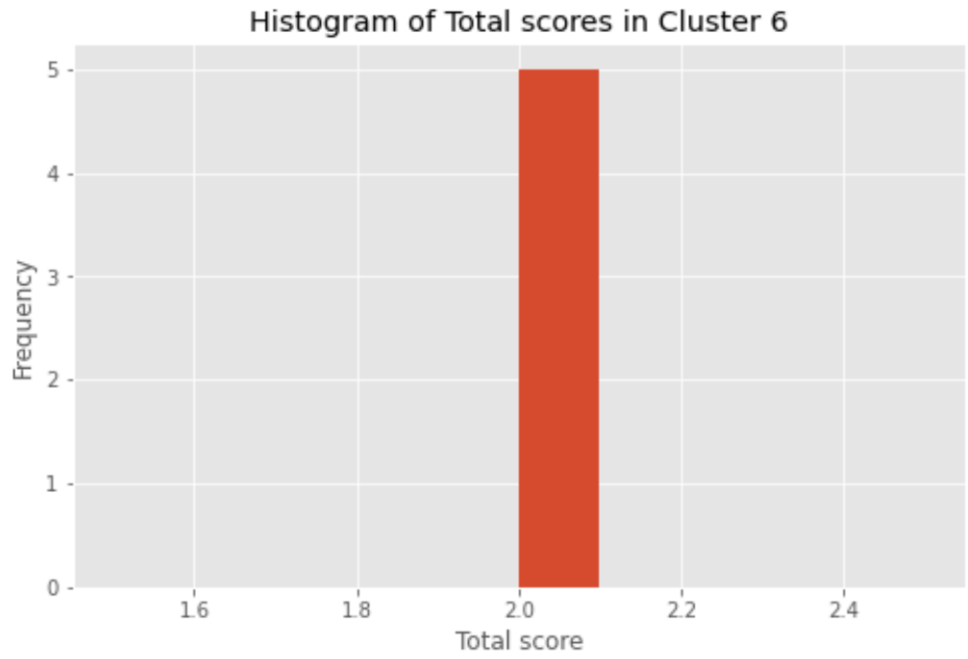




Crime rate > average, # Restaurants/Shopping places < average with Total score = 2

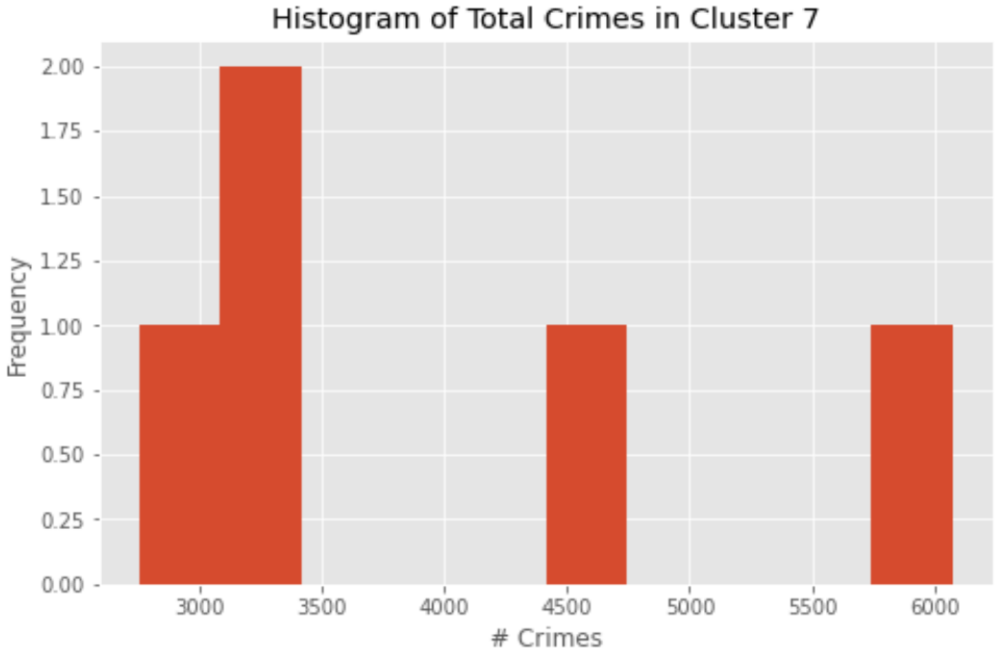
K MEANS CLUSTERING- CLUSTER 6





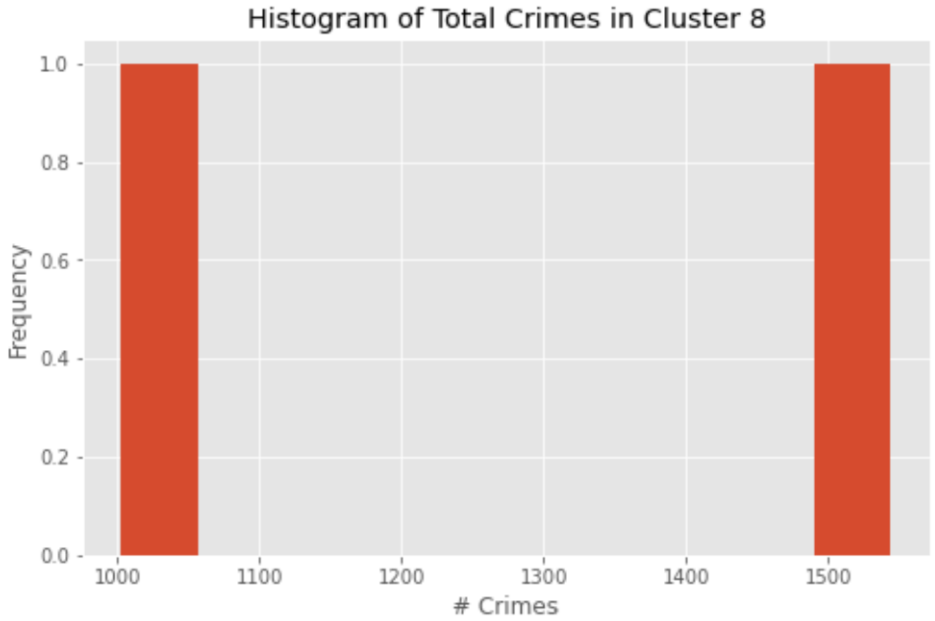
Crime rate < average, # Restaurants/Shopping places < average with Total score = 2
All neighborhoods have crime incidents less than 1000 in this cluster

K MEANS CLUSTERING- CLUSTER 7



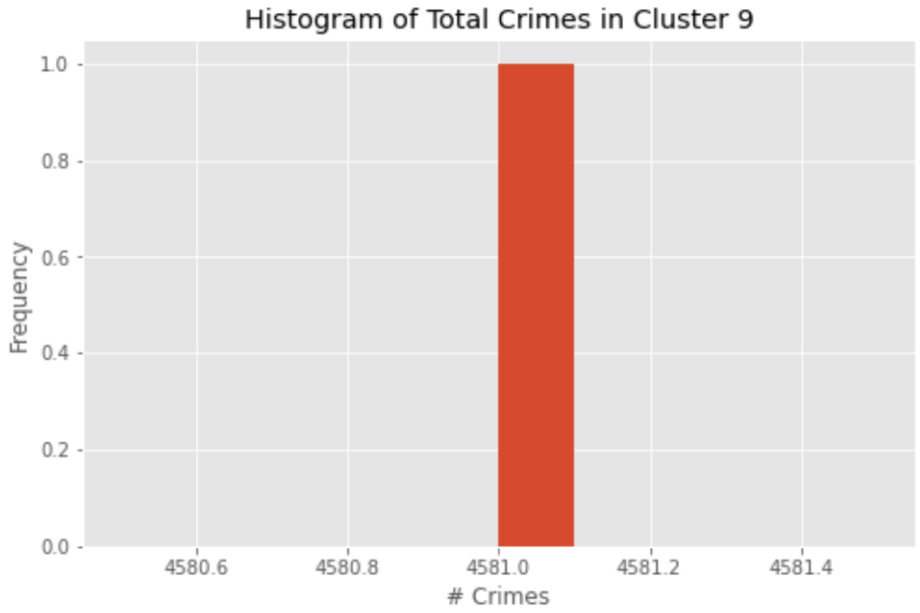
Crime rate > average, # Restaurants/Shopping places < average with Total score = 1

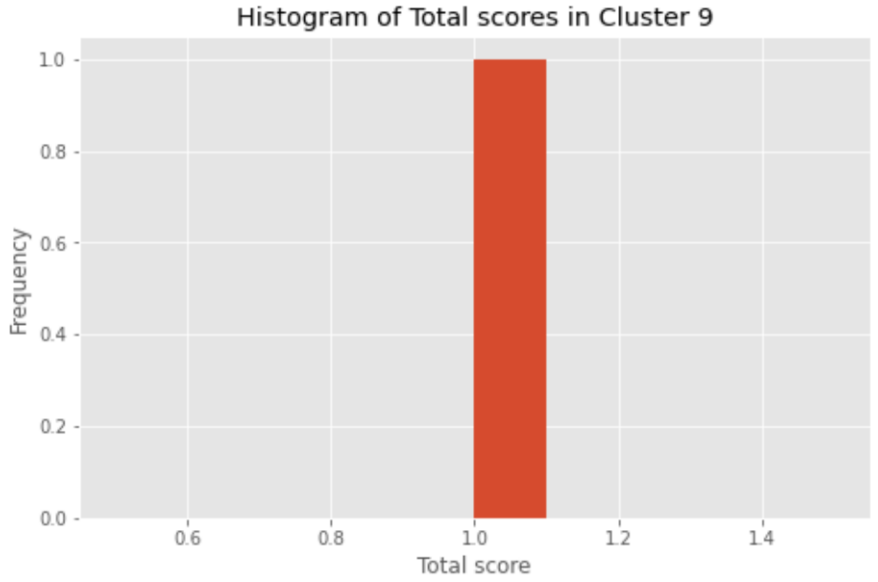
K MEANS CLUSTERING- CLUSTER 8



Crime rate < average, # Restaurants/Shopping places < average with Total score = 1

K MEANS CLUSTERING- CLUSTER 9

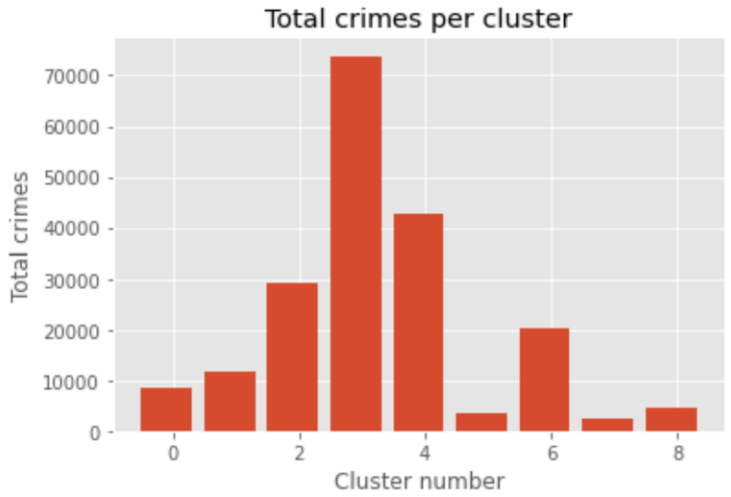




Crime rate > average, # Restaurants/Shopping places < average with Total score = 1

NEIGHBORHOODS TOTAL CRIMES/TOTAL SCORE

Considering cluster label index = 0 to 8



Cluster 3 – Crime incidents > 70000
Cluster 5,7 – less than 10000 crimes

Cluster 3 – Total score =0
Cluster 1 – Total score = 3

CONCLUSION

- Best option – Of all the clusters from 1 to 9, cluster 2 with less crimes than average, total score = 3, enough shopping/eatery places.
- Least desirable neighborhood - Of all the clusters from 1 to 9, cluster 4 exceeds 70000 crimes in total and total score =0 , not enough shopping / restaurant places.
- More factors can be added to the neighborhood selection for future purposes.