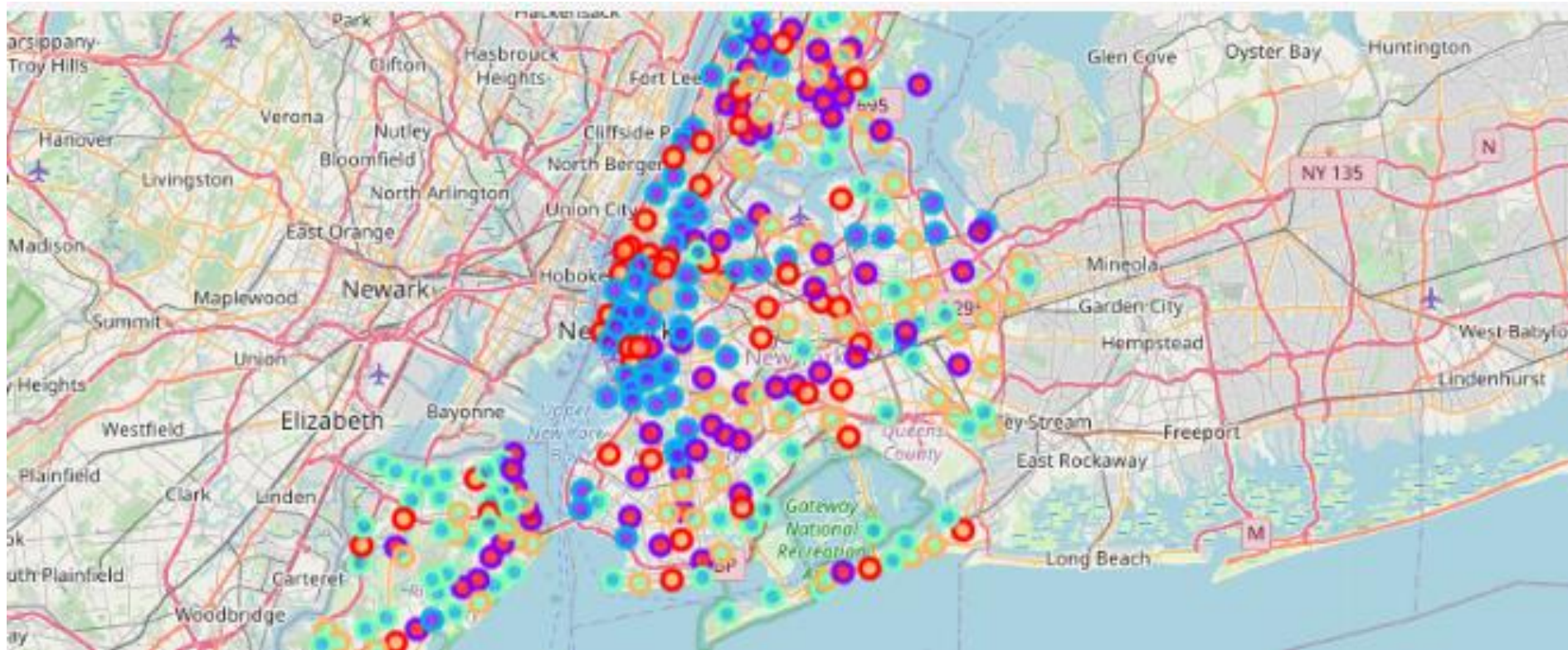# Real State App – Helping hand

Overview

❑ App for real estate agencies

➢ Allows we to compare neighborhoods in different cities (New York - Toronto)

➢ Delivering the information in table and map format

➢ Categorizing neighborhoods in three dimensions

- Economic
- Social
- Services

➢ Dimensions Populated from Foursquare

Data acquisition

❑ New York/Toronto Venues Grouped

➢ Venues are grouped by neighborhoods

➢ Making a sum in the corresponding dimension.

New York

	Unnamed: 0	Neighborhood	ECO	SER	SOC
0	0	Allerton	25	4	0
1	1	Annadale	11	1	0
2	2	Arden Heights	3	1	0
3	3	Arlington	3	3	0
4	4	Arrochar	19	4	0
5	5	Arverne	14	7	0
6	6	Astoria	48	0	2
7	7	Astoria Heights	7	4	1
8	8	Auburndale	18	2	0
9	9	Bath Beach	47	3	0

Toronto

	Unnamed: 0	Neighborhood	ECO	SER	SOC
0	0	Agincourt	4	0	0
1	1	Alderwood, Long Branch	7	1	1
2	2	Bathurst Manor, Wilson Heights, Downsview North	22	1	0
3	3	Bayview Village	4	0	0
4	4	Bedford Park, Lawrence Manor East	24	0	0

Clustering Algorithm
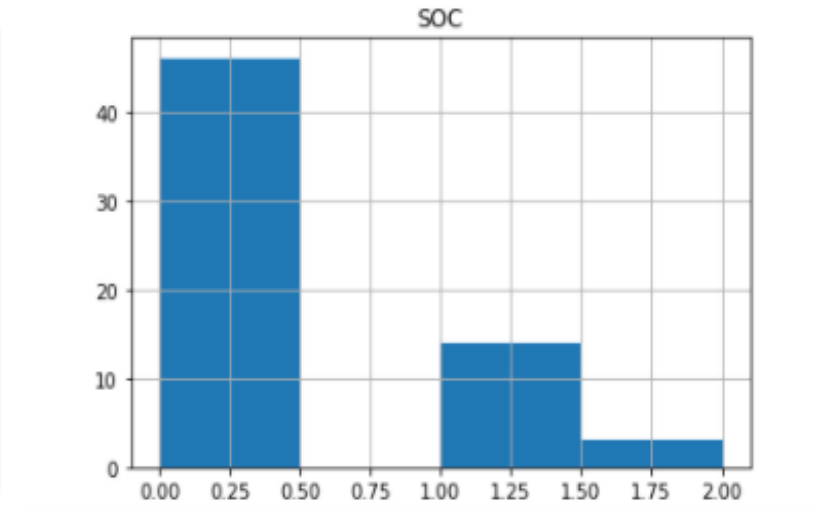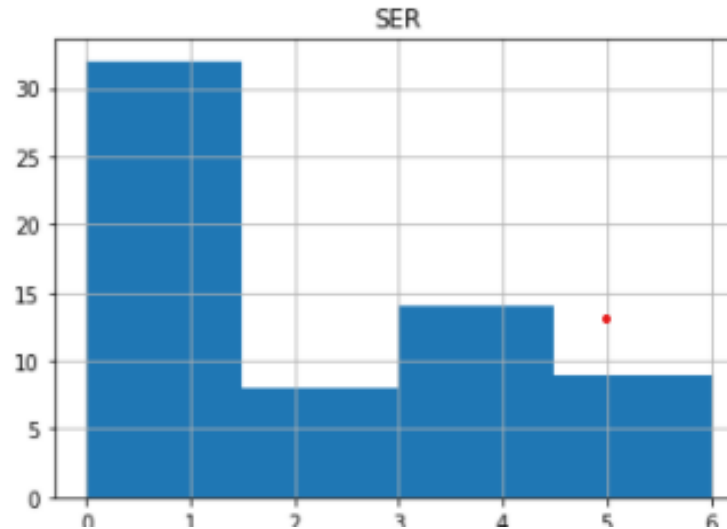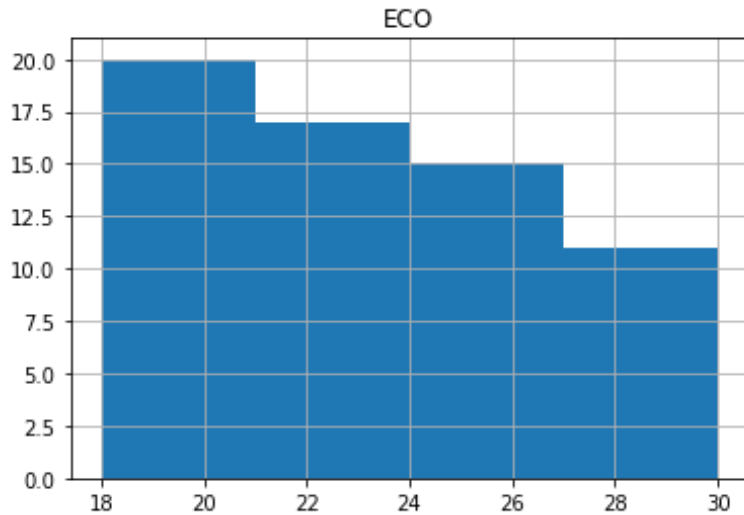
➢   Use Kmeans with k =5 (outcome five clusters)

➢   Using dataframe New York venues grouped

➢   Making a dataframe to analyze.

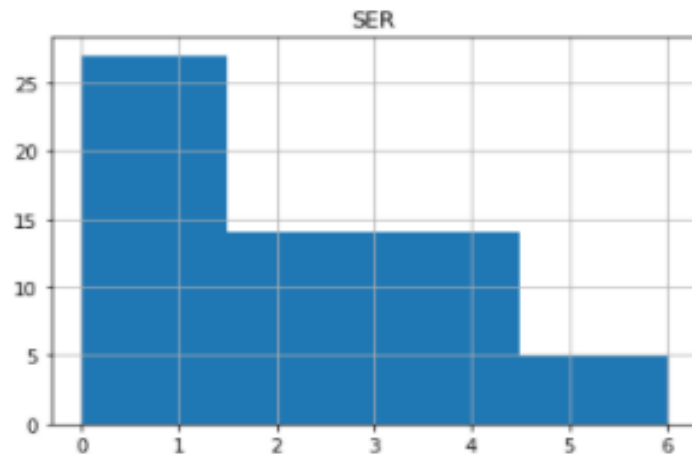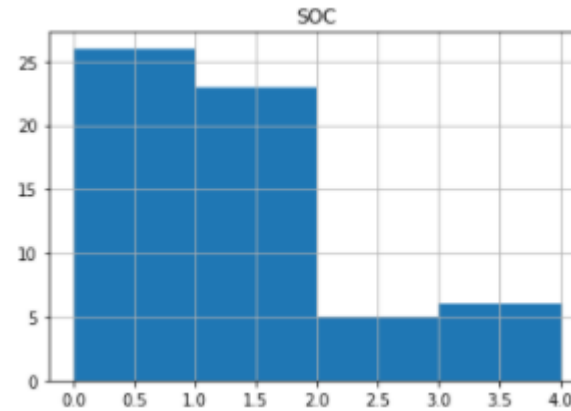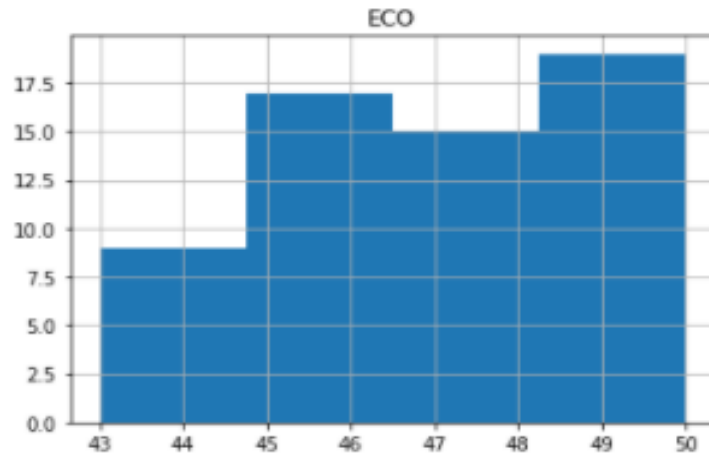➢   Printing in a map to analyze:
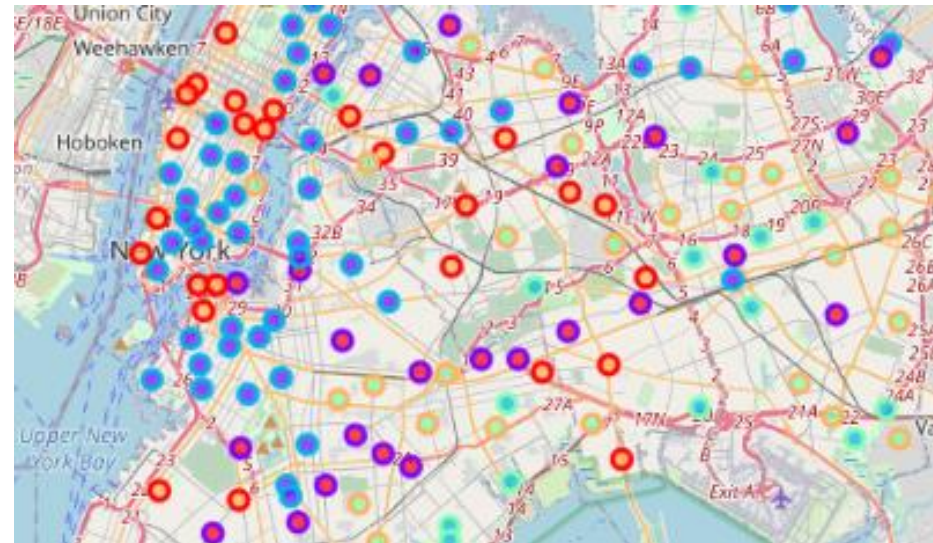
Clustering Algorithm

❑   Cluster 0   (Top 3)

➢ Sufficient public and private services
➢ Acceptable infrastructures
➢ Enough shops at street level.

Clustering Algorithm
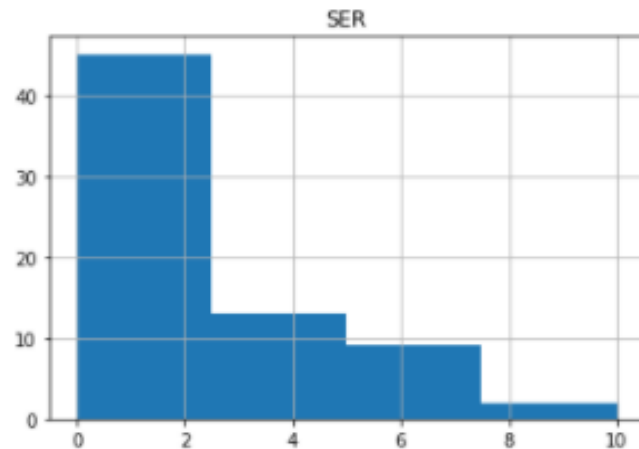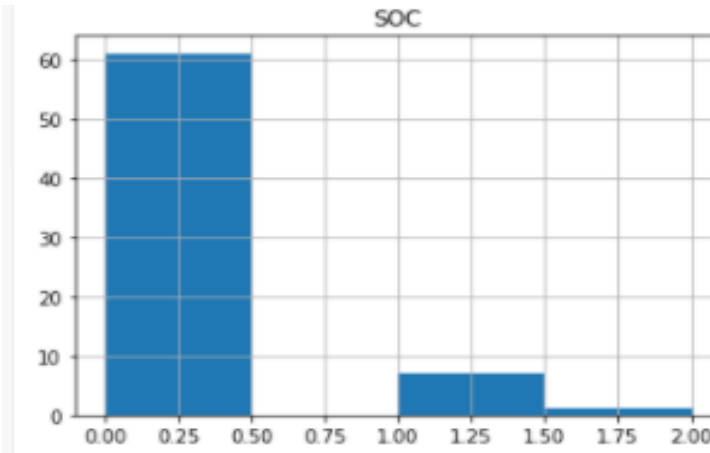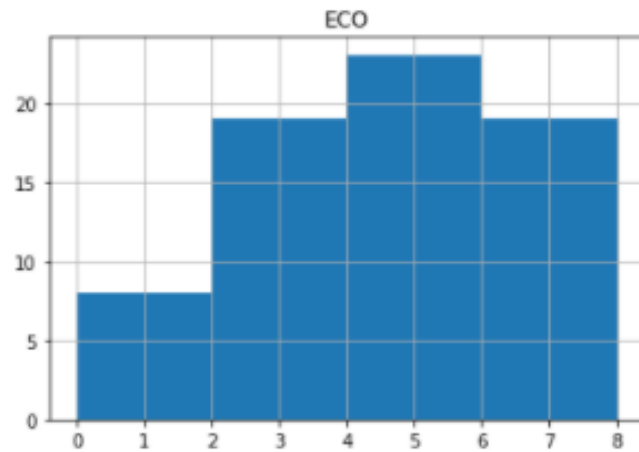
❑ Cluster 1 (Top 1)

ECO

SOC

SER

➢ The best public and private services
➢ Excellent infrastructures
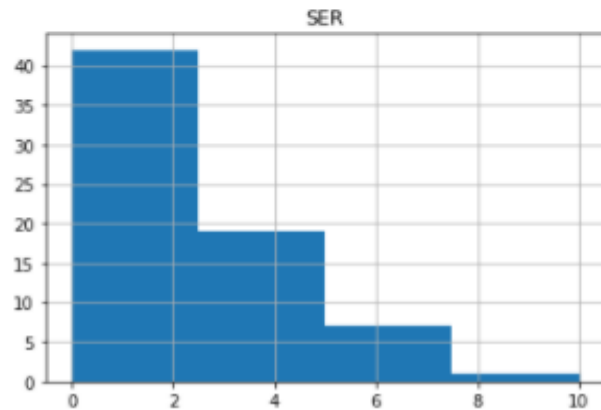➢ Many shops at street level.
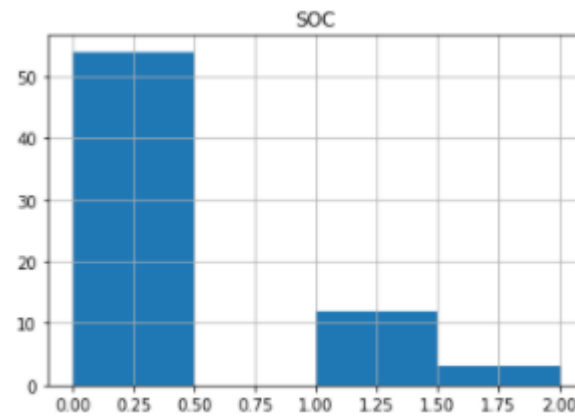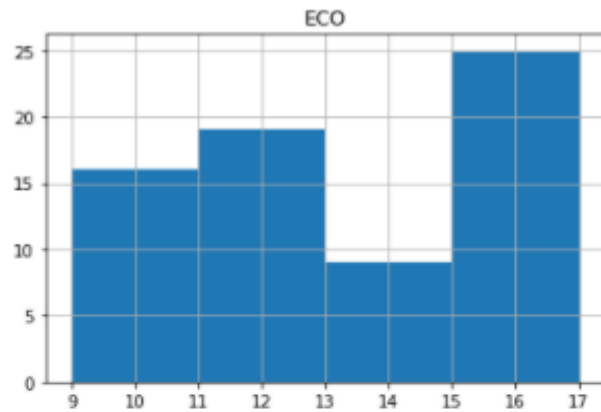
Clustering Algorithm

❑ Cluster 2   (Top 5)

- ➤ Few public and private services
- ➤ Little infrastructure
- ➤ Few businesses at street level
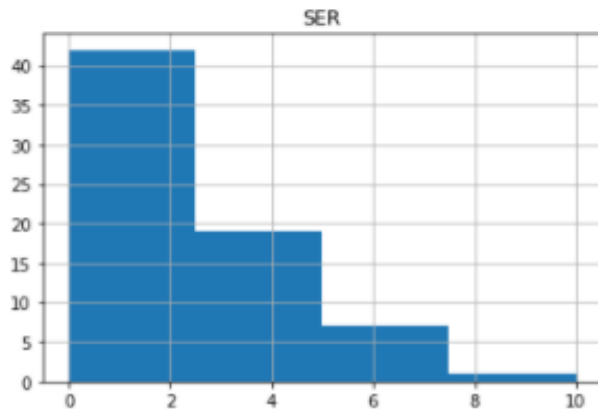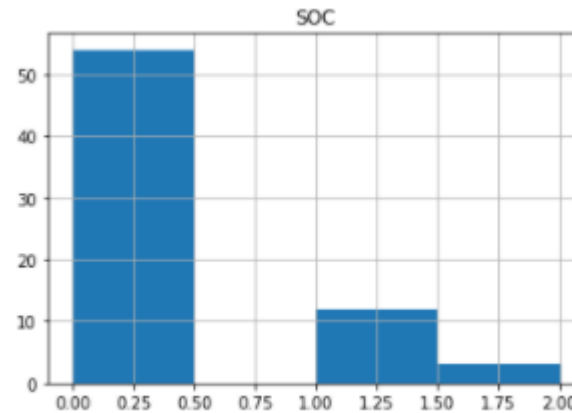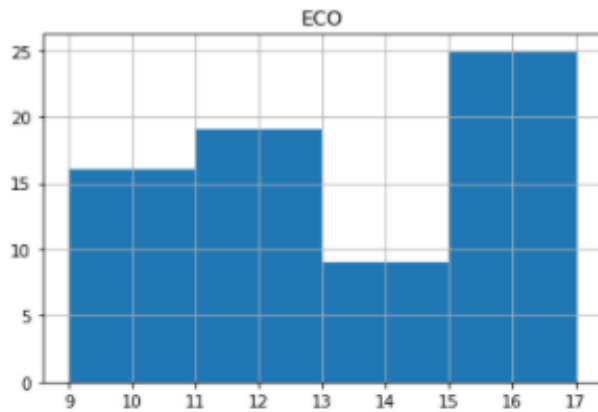
Clustering Algorithm
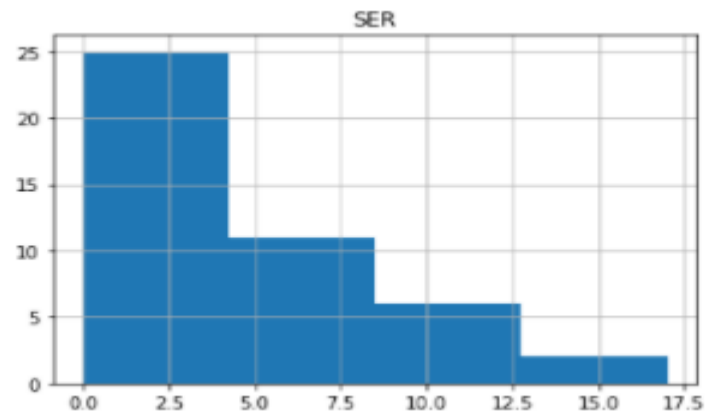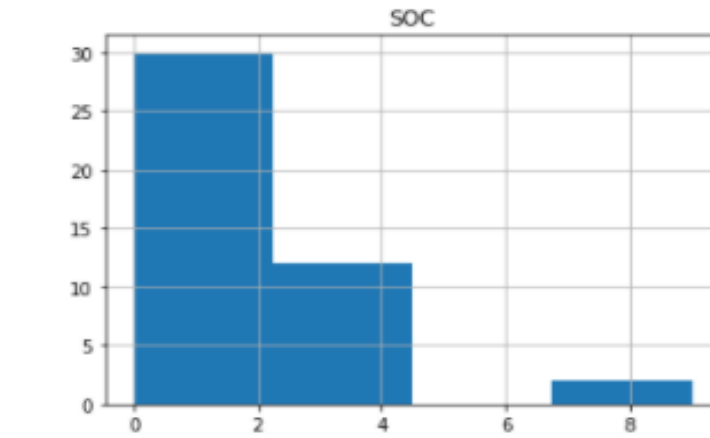
❑ Cluster 3 (Top 4)

➢ Few public and private services
➢ Little infrastructure a
➢ Few shops at street level.

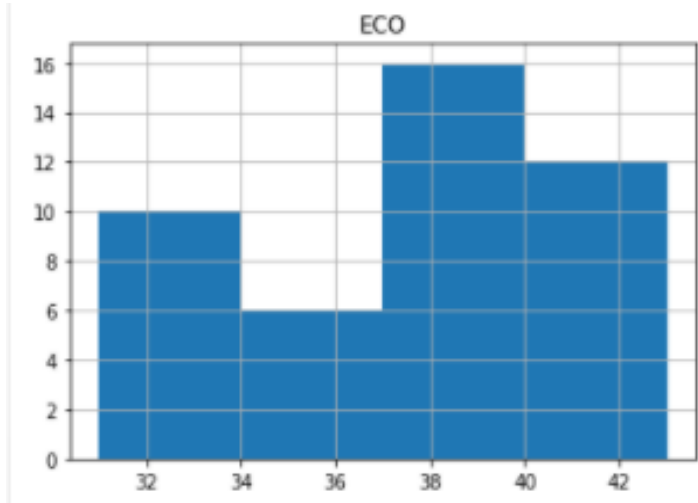# Clustering Algorithm

❑ Cluster 4 (Top 4)

> ➢ Few public and private services
> ➢ Little infrastructure a
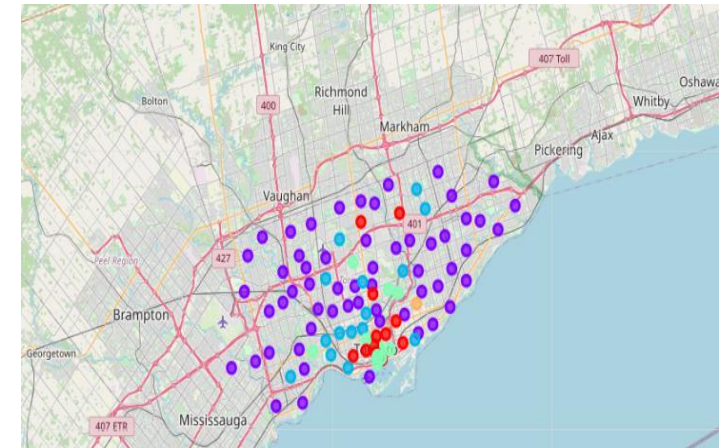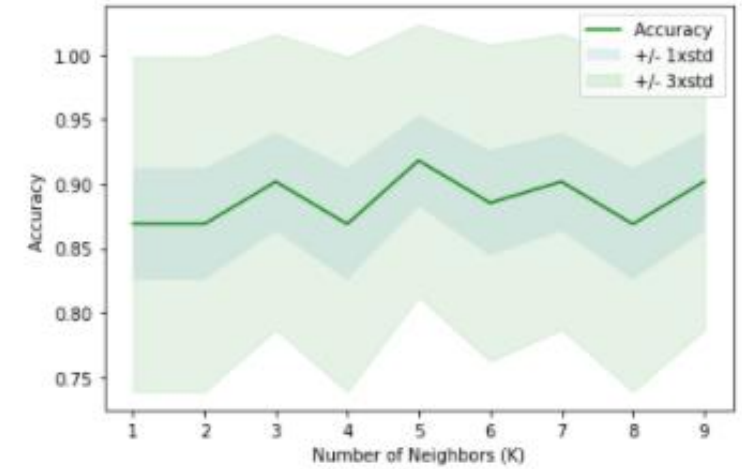> ➢ Few shops at street level.

Clustering Algorithm

❑ Cluster 5 (Top 2)

> High public and private services
> Abundant infrastructure
> Many shops at street level

Classification algorithm

➢ Use k-nearest neighbors algorithm with ks =5

➢ Accuracy = 0,91

➢ Printing in a map to analyze:

Conclusions

Our application helps us to homogenize the comparison of neighborhoods in different cities, both visually and with the two dataframes generated. (NewYork_print, Toronto_print).

It can be applied to many use cases. We are going to carry out a practical case so that it is better understood.

Client lives in the "North Riverdale" neighborhood and wants to see Toronto neighborhoods similar to this one. As it belongs to Cluster 0 we have to look for a neighborhood in Toronto inferenced with Cluster 0.

Future directions

To improve the state of the art, progress could be made in several directions.

One of them would be to incorporate other data sources that allow us to add more dimensions, for example subjective, which are those that are based on the opinion that neighborhoods have about their neighborhood. Dimensions that take into account the social circumstances of the neighbors, unemployment, drugs, etc. could also be included.

Other algorithms could also be checked to see if they improve accuracy.

Finally, more cities could be included and thus expand the catalog of cities to be compared.