Capstone Project - The Battle of Neighborhoods

Capstone Project - The Battle of Neighborhoods

- Normally there are so many things that are different from one country to another or one city to another.
- But there are also some things based on what we can say some places are similar to other.
- For that reason we want to cluster the neighborhoods of the New York city and the city of Toronto.

Data acquisition and preprocessing

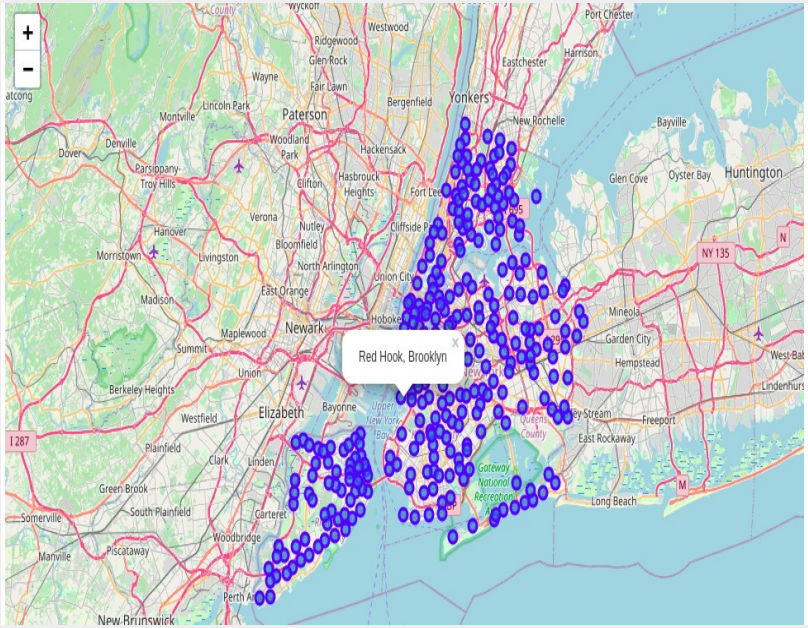
- Data for the New York city:

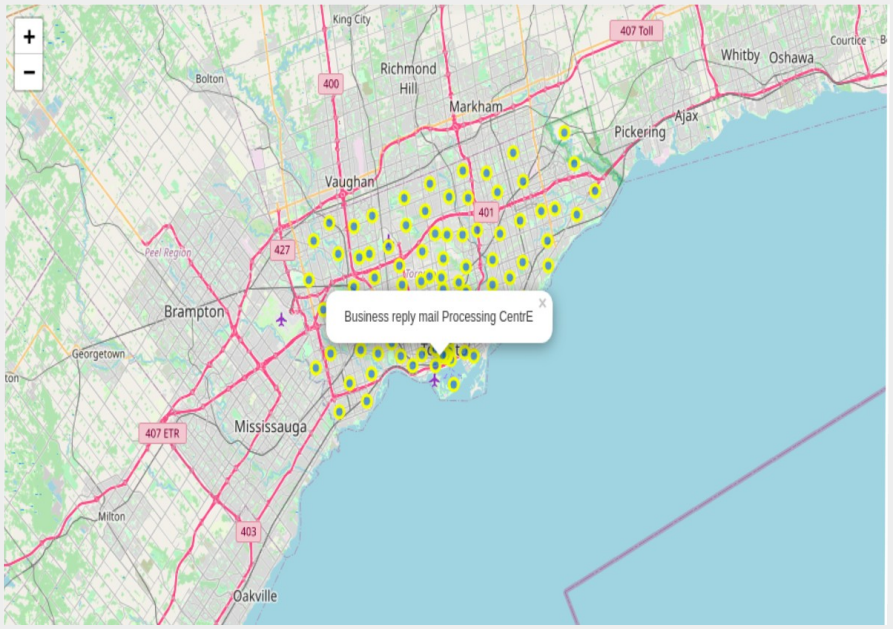
  link: https://cocl.us/new_york_dataset

- Data for the city of Toronto:

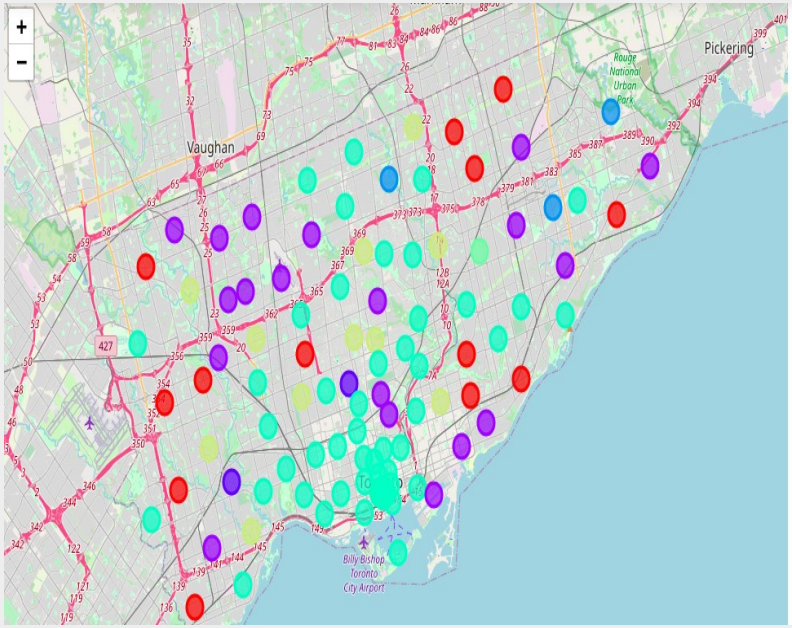
  link: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

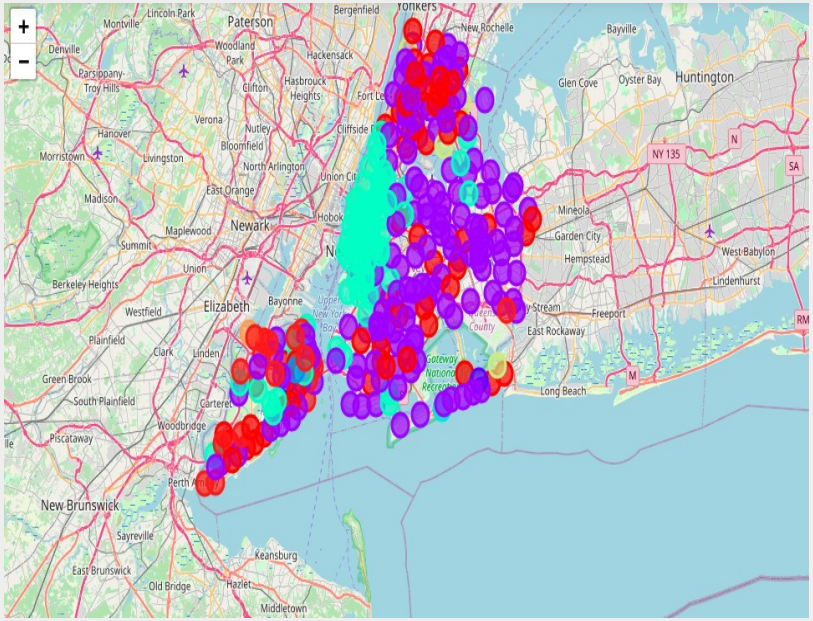
Visualizing the data before clustering



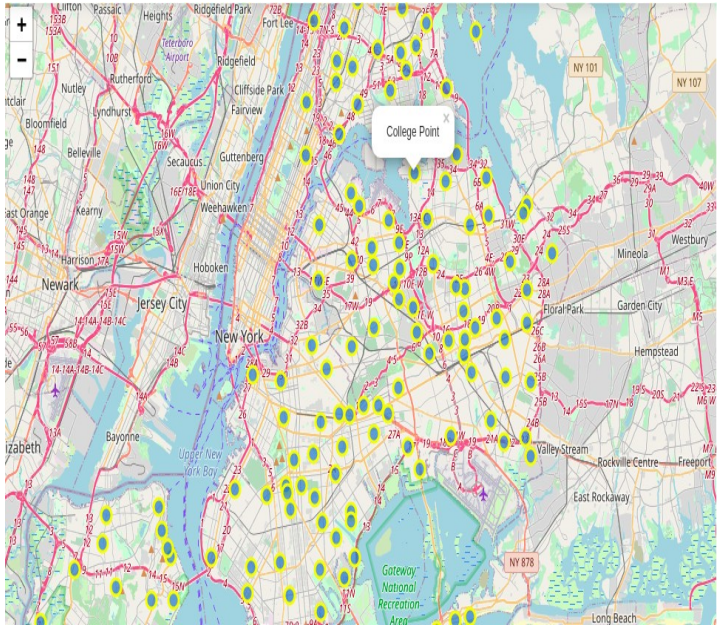


Visualizing data after clustering

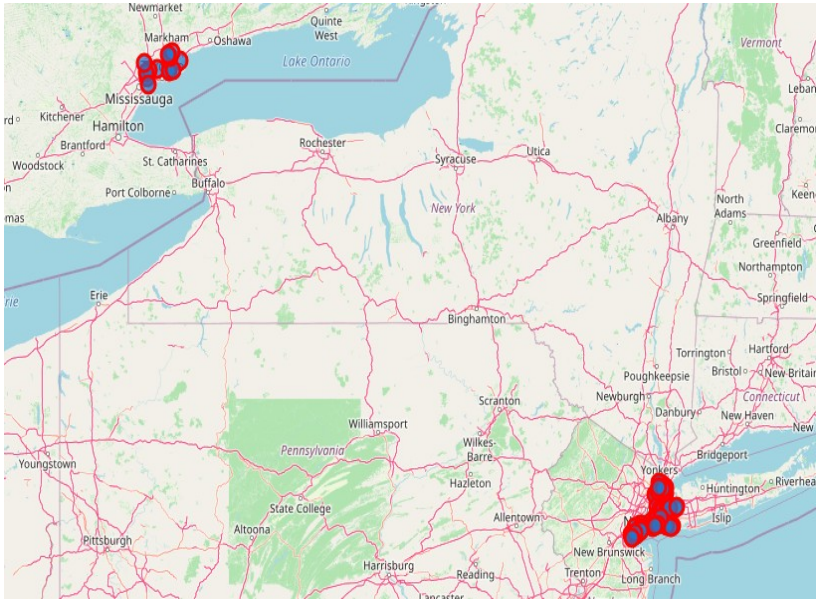




Examples of some clusters

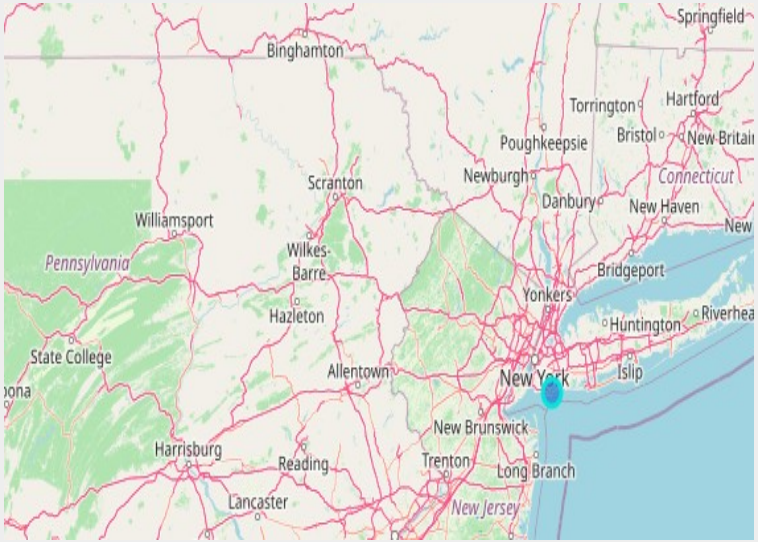


Cluster 1

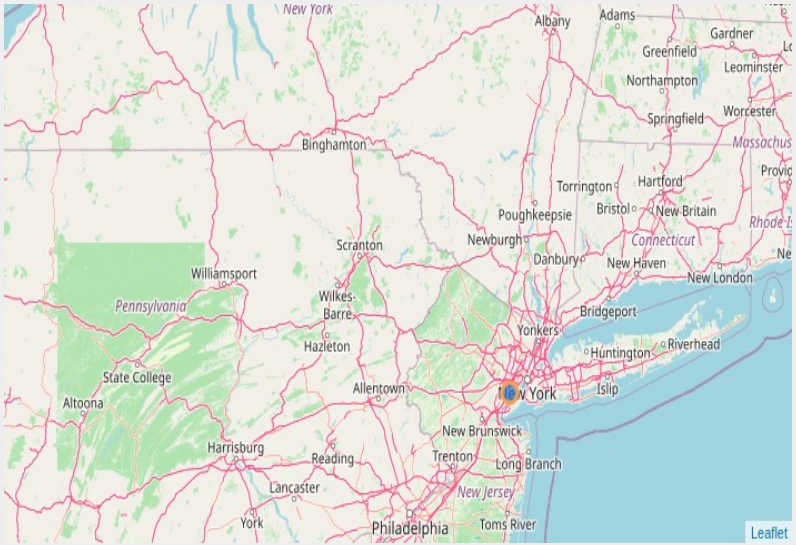


Cluster 10

Cluster with single unique value



Cluster 4



Cluster 8

Results & Discussion

- Total number of 10 clusters are there, and each cluster consists some values
- In cluster number 1,2,3,5,7,9 & 10 consist of values which are on both cities
- Values in the same clusters are similar places
- 3 clusters have only one item

Conclusion

- Built a model by clustering the neighborhoods of two cities
- Items in the same clusters are similar categorically
- Could be achieved more impactful result by changing the cluster numbers
- This model can be used to predict different values by using different machine learning algorithms

THANK YOU