The Battle of Neighborhoods - Coursera Capstone Project

Finding an optimized method for delivering fresh produce in the nearby restaurants/cafe in North York, Toronto

IBM Course 9 – Applied Capstone Project

By: Shaifali

Problem Formulation

A farmer wants to expand its business in the local neighborhood and wants to deliver fresh produce everyday in the local areas of North York, Toronto. There should be timely delivery of fresh produce, veggies, bread, coffee etc. in order to successfully run the business. In order to achieve that, there should be some optimal method to deliver in every single restaurant/coffee shops/cafe nearby with no delay at all to any place. There are 10 delivery trucks available to use and all the areas to be covered. The best way would be to divide the potential merchants into segment based on their locations and address the mechanism separately for each segment of merchants. This would allow the farmer to use all the resources in a efficient manner and hence help rum a successful business plan.

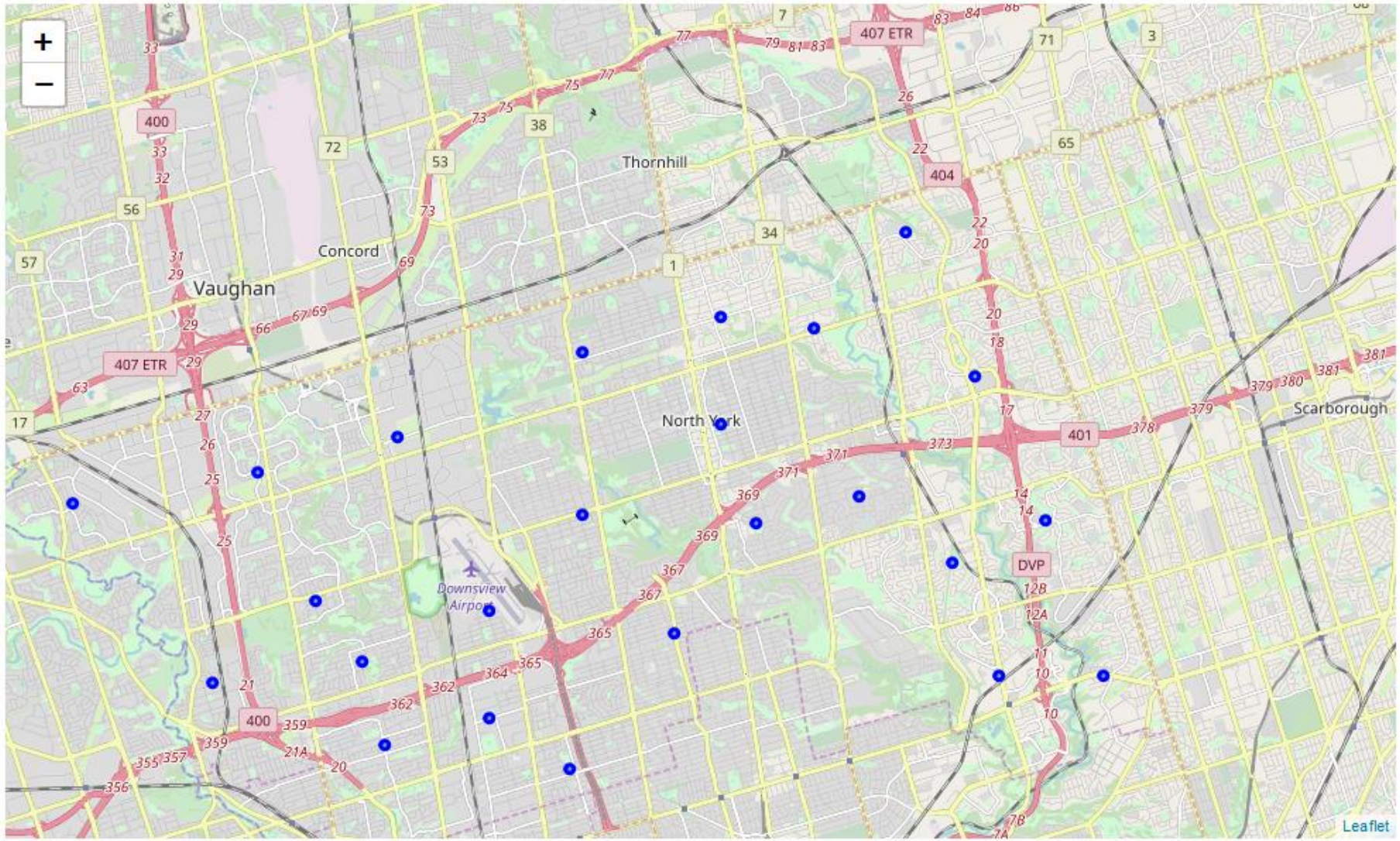
Python Packages and Dependencies

- Pandas - Data Analysis Library
- NumPy – Library to handle data as a DataFrame
- JSON – Library to handle JSON files
- Geopy – Used to find geospatial coordinates
- Requests – Library to handle https requests
- Matplotlib – Used to generate plots in python
- Sklearn – Library to implement Machine Learning techniques
- Folium – Map rendering library

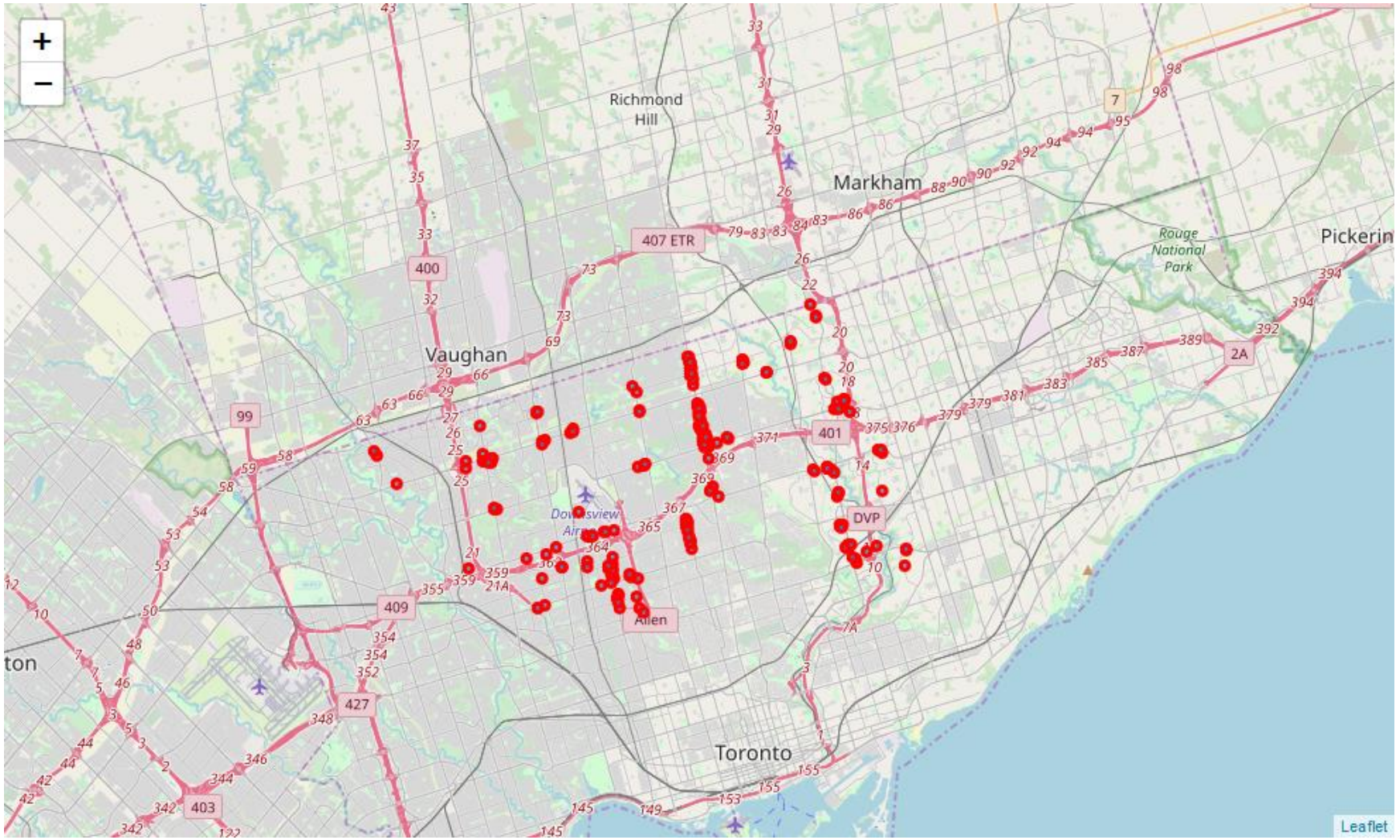
Sample of Data for the neighborhood North York, Toronto

	Postcode	Borough	Neighbourhood	Latitude	Longitude
17	M2H	North York	Hillcrest Village	43.803762	-79.363452
18	M2J	North York	Oriole,Fairview,Henry Farm	43.778517	-79.346556
19	M2K	North York	Bayview Village	43.786947	-79.385975
20	M2L	North York	York Mills,Silver Hills	43.757490	-79.374714
21	M2M	North York	Newtonbrook,Willowdale	43.789053	-79.408493

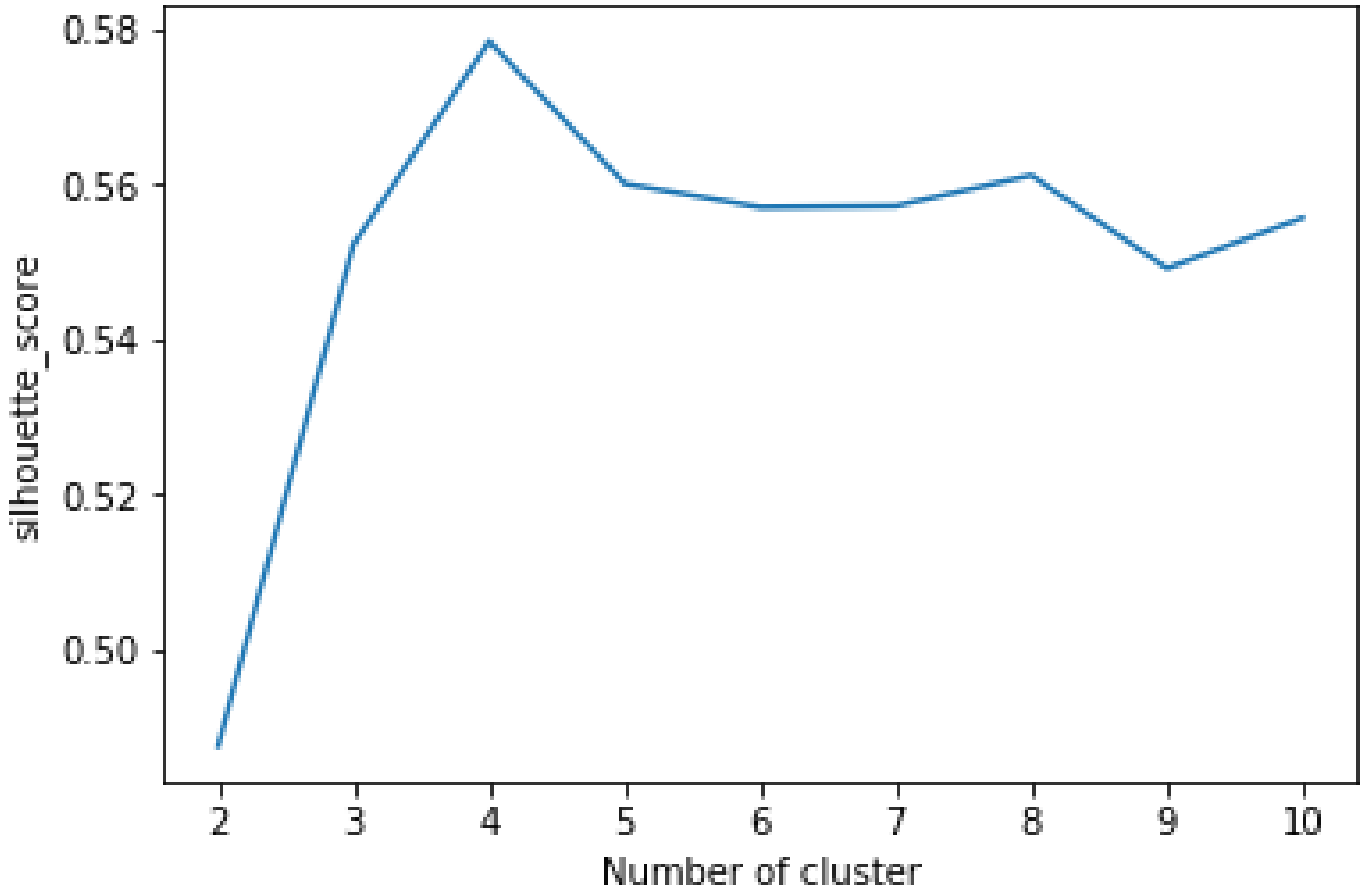
Neighborhood Distribution



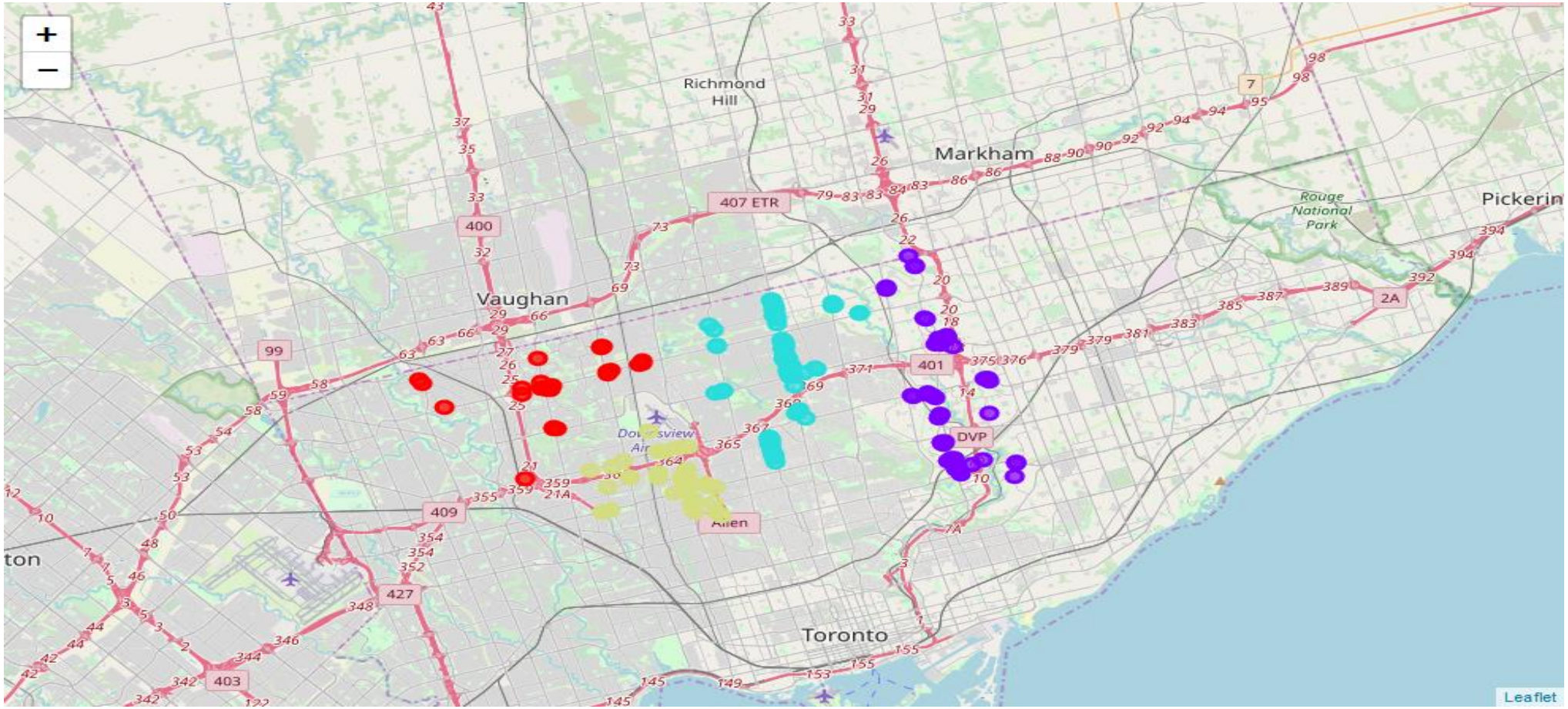
Venue Distribution



Results – based on K value and Silhouette Coefficient



K-means Clustering (K = 4)



Discussion and Conclusion

The techniques and analysis used in the project shows that the optimal way for the farmer to use the delivery trucks in order to cover every interested venue and deliver the produce efficiently can be performed using the clustering method. The efficient number of clusters lies between K = 4 to K = 6. Here, we have suggested K = 4 as the cluster number in order to perform the task in an optimal way which means there will be 4 delivery trucks in total to be used for delivery.