Data Science Capstone Project

For the Completion of IBM
Data Science Professional
Certificate

Introduction/Business Problem:

- Japanese cuisine is one of the most famous cuisine in the world. This is because of they use only the best ingredients for their food. A client of mine is planning to supply some ingredients to Japanese restaurant that are located at Downtown Toronto. With this, our goal is to generate clusters of these restaurant for better service and deliveries.

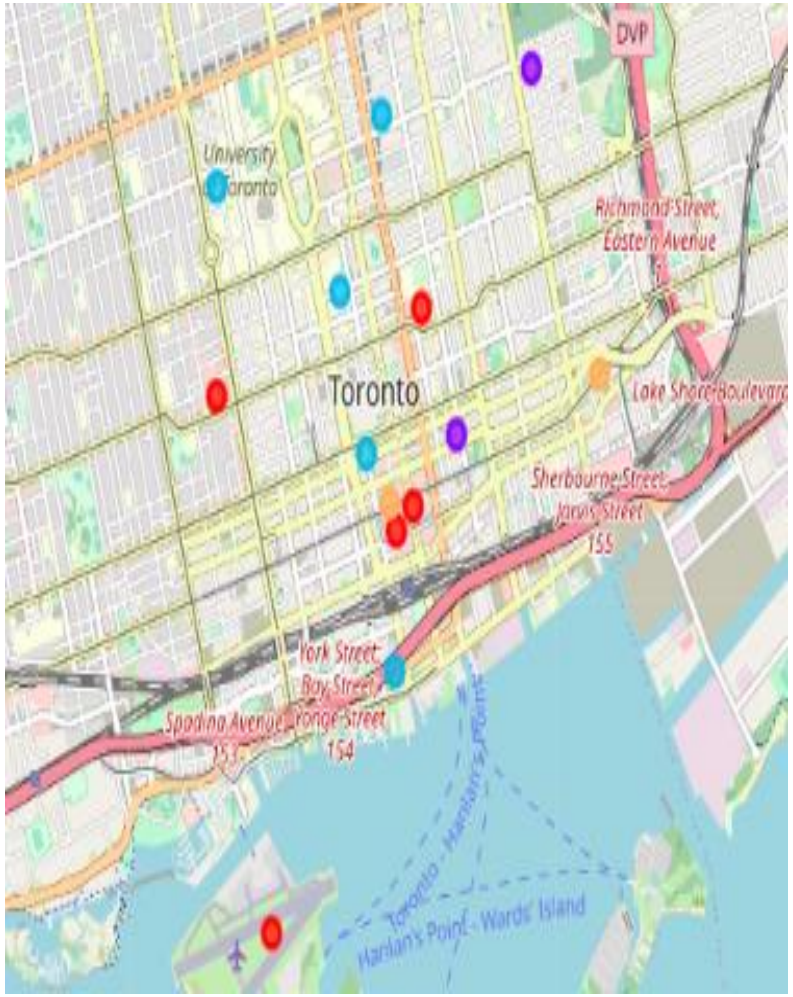
Description of Data

- As instructed by the assigned task, the data that will be used will come mainly from foursquare. Foursquare is an American Technological Company from New York focusing on location data. We will use their product, foursquare data, to get the location data about the Japanese restaurants in Toronto.
- The data is not only limited to Foursquare data, we can also use other data. For this project we will also consider data coming from Wikipedia. Using beautifulsoup, we can scrape the data Neighborhood and Boroughs in Toronto. We will then merge the Wikipedia data with the COCL data which has the coordinates of the Neighborhood in Toronto.

Methodology

- Apply Webscraping using Beautifulsoup to create dataframe.
- Apply for Foursquare credentials.
- Extract necessary data through Foursquare database.
- Perform One Hot Coding
- Apply K-means Clustering
- Create Map using Folio
- Print Clusters no. 1 to 5

Results



Below are the 5 clusters or 5 routes for delivery:

Cluster 5 (Delivery Route 1):

	Postcode	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	Total Restaurants	Japanese Restaurants
47	M5A	Downtown Toronto	Regent Park	43.654260	-79.360636	4	12	3
60	M5X	Downtown Toronto	First Canadian Place	43.648429	-79.382280	4	12	2

Cluster 4 (Delivery Route 2):

	Postcode	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	Total Restaurants	Japanese Restaurants
44	M4W	Downtown Toronto	Rosedale	43.679563	-79.377529	3	27	3

Cluster 3 (Delivery Route 3):

	Postcode	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	Total Restaurants	Japanese Restaurants
46	M4Y	Downtown Toronto	Church and Wellesley	43.665860	-79.383160	2	17	2
50	M5G	Downtown Toronto	Bay Street	43.657952	-79.387383	2	17	3
51	M5H	Downtown Toronto	Richmond	43.650571	-79.384568	2	18	4
52	M5J	Downtown Toronto	Harbourfront	43.640816	-79.381752	2	20	3
57	M5S	Downtown Toronto	University of Toronto	43.662696	-79.400049	2	17	2

Cluster 2 (Delivery Route 4):

	Postcode	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	Total Restaurants	Japanese Restaurants
45	M4X	Downtown Toronto	St. James Town	43.667967	-79.367675	1	32	5
49	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418	1	32	5

Cluster 1 (Delivery Route 5):

	Postcode	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	Total Restaurants	Japanese Restaurants
48	M5B	Downtown Toronto	Garden District	43.657162	-79.378937	0	14	3
53	M5K	Downtown Toronto	Toronto Dominion Centre	43.647177	-79.381576	0	15	2
54	M5L	Downtown Toronto	Commerce Court	43.648198	-79.379817	0	14	2
58	M5T	Downtown Toronto	Kensington Market	43.653206	-79.400049	0	15	1
59	M5V	Downtown Toronto	CN Tower	43.628947	-79.394420	0	15	1

Conclusion

- Overall, we achieved the main goal of this capstone project. That is to provide delivery route through clustering for a client.