Finding the best locations for expansion for the Example Franchise in Manhattan and Toronto.

Shawn Lim Junhe

28 May 2020

1. Introduction and Problem

Example Pte Ltd is a Singaporean Franchise looking to expand its operations overseas after a successful launch in Singapore and subsequent expansion into South East Asia. The strategic management of the company has decided on New York City, Manhattan and Toronto, Canada as possible locations to spearhead its expansion into the North American region. However, the team is currently unsure about which city would be most suitable for this expansion. This choice of country is of paramount importance as it sets the tone and perception of the company in the North America Region and is likely to make or break the success of subsequent expansion. Due to the importance of this choice, the company has engaged its data science team to suggest which city and locations would be the most suitable for Example's entrance into the North America Region. Analysis from previous expansion has shown that locations that are the most similar to Example's most successful store in Orchard, Singapore tend to also enjoy a high degree of success as well. As such the team has decided to select locations that are the most similar to the Orchard Store.

2. Data Acquisition and Cleaning

2.1 Data Sources

Postal codes of the neighborhoods in Toronto was scraped from Wikipedia and their latitudes and longitudes were obtained using their respective postal codes from here. The latitudes and longitudes of the Orchard stall

was estimated to the in the vicinity of the Orchard Mass Rapid Transport (MRT) station which was obtained from here. The data for Manhattan was obtained from here. By obtaining the latitudes and longitudes of the various neighborhoods, data about venues within their vicinity could then be obtained through the Foursquare API.

2.2 Data Cleaning

2.2.1 Toronto

The data scraped from the Wikipedia page had postal codes that with Boroughs that were "Not assigned" which had to be dropped. The geospatial data could then be merged using the postal codes. Neighborhoods that were outside Toronto was then removed and neighbors with the same postal codes were combined. The geospatial data for the Orchard store was then added to the data set and location information from Foursquare was obtained through the Foursquare API. After the information about the venues for each neighborhood was gathered, they could then be encoded into binary values and grouped so that the frequency of each venue category for each neighborhood could be obtained. The venue categories were then the features on which classification could be performed on. The foursquare data for the orchard store was added to both datasets so that clustering could be done.

2.2.2 Manhattan

The cleaning process for the Manhattan data was similar to the Toronto data, except that the latitudes and longitudes for the Manhattan neighborhood were already provided.

2.3 Feature Selection

After data cleaning, there were 1712 samples and 245 features in the Toronto data and 3177 samples and 331 features for the Manhattan data. As the data for the orchard store was included in both datasets, the

features of the orchard store was present in both datasets and should not bias the results. None of the features were dropped as the goal of this project was to find out the which neighborhoods were most similar to the orchard store.

3. Exploratory Data Analysis

```
----Orchard----
                venue  freq
0             Boutique  0.10
1     Sushi Restaurant  0.07
2              Bakery  0.06
3               Hotel  0.06
4  Japanese Restaurant  0.05
```

Table 1

Neighborhood	Orchard
1st Most Common Venue	Boutique
2nd Most Common Venue	Sushi Restaurant
3rd Most Common Venue	Hotel
4th Most Common Venue	Bakery
5th Most Common Venue	Japanese Restaurant
6th Most Common Venue	Shopping Mall
7th Most Common Venue	Cosmetics Shop
8th Most Common Venue	Chinese Restaurant
9th Most Common Venue	Indonesian Restaurant
10th Most Common Venue	Bubble Tea Shop

Table 2

The table 1 shows the frequencies of the 5 most common venues within 500 meters of the Orchard store and table 2 shows the top 10 most common venues in descending order. The most common venue category was Boutique stores which might explain the store's success as it is situated in a higher end

part of the city. Further details about the frequencies of the different venues for neighborhoods in the respective cities can be found in the code.

4. Predictive Modeling

4.1 Classification models

K – means clustering was chosen to classified the different neighborhoods based on the frequencies of the most common venues. Classification based on the similarity of venues in the different neighborhoods was chosen since as I have mentioned earlier, previous analysis of successful expansions have shown that stores that were similar to the Orchard store had the highest degree of success. The number of neighborhoods in each of the cities were about the same was about the same (around 40) and the number of features for the Toronto neighborhoods was 241 and the number of features for the New York neighborhoods was 331.

4.2 Choosing the optimal k

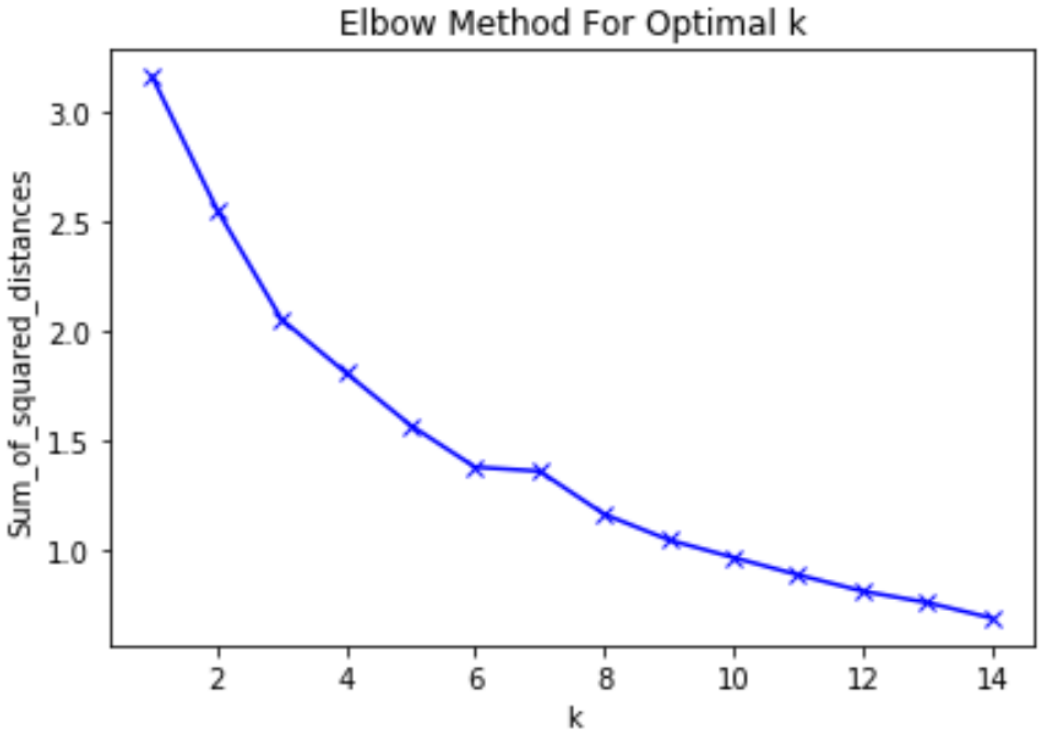

Figure 1: Optimal k for Toronto Dataset

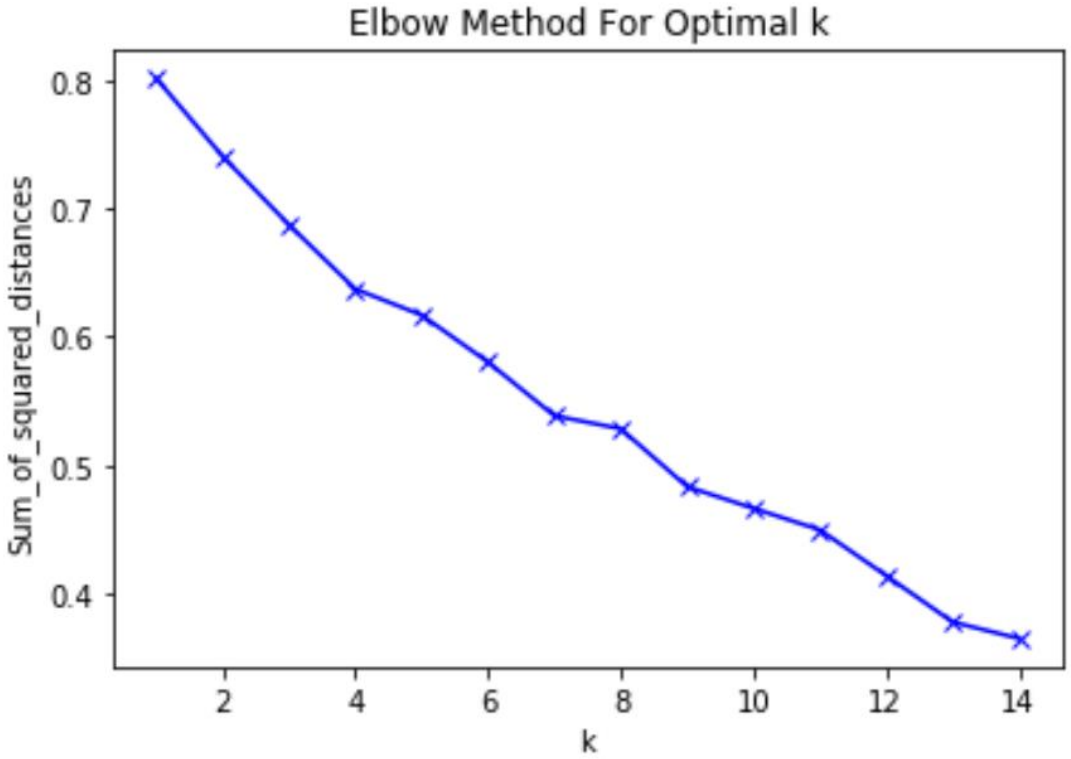

Figure 2: Optimal k for Manhattan Dataset

I chose the optimal k clusters for k-means clustering using the elbow-method for both k-means clustering of the different neighborhoods. Using the method, I found that the optimal k for the Toronto dataset was 6 and the optimal k for the Manhattan dataset was 4. The optimal k for the Manhattan dataset was harder to determine as there was no clear elbow point, so I took the first elbow point which was k equals to 4.

4.3 Results

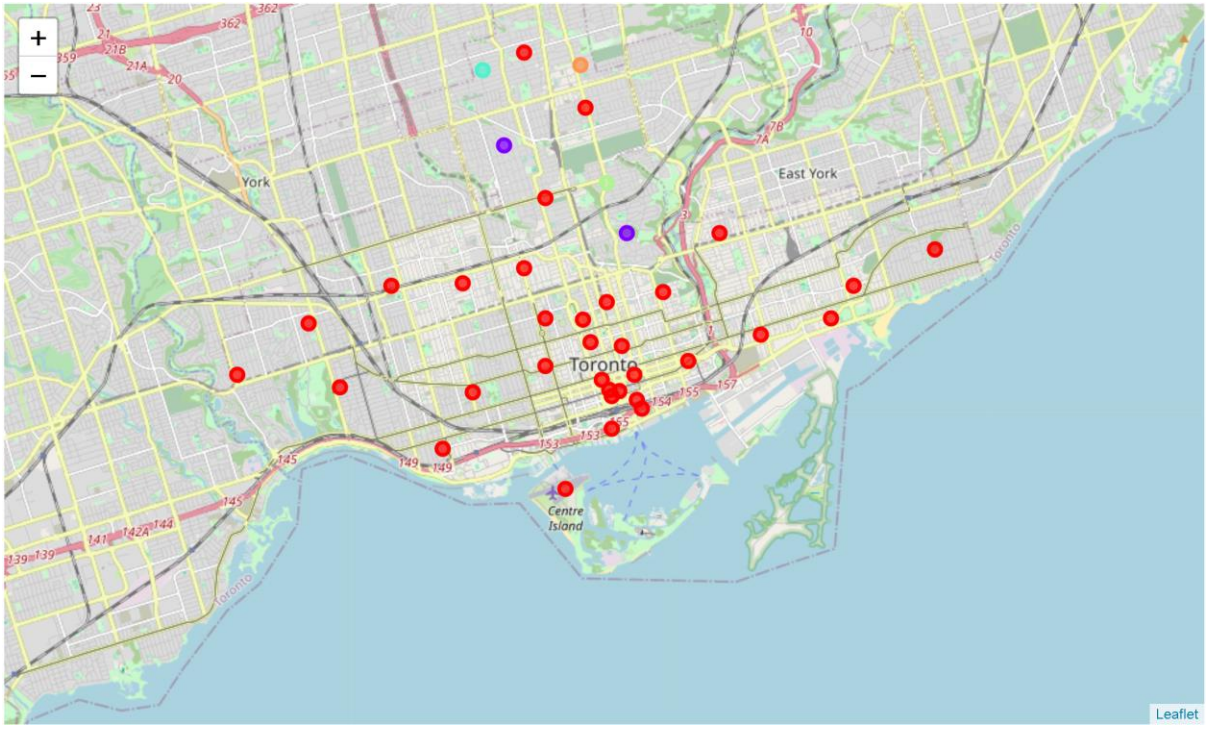

Figure 3: Clustering results for the neighbourhoods of Toronto

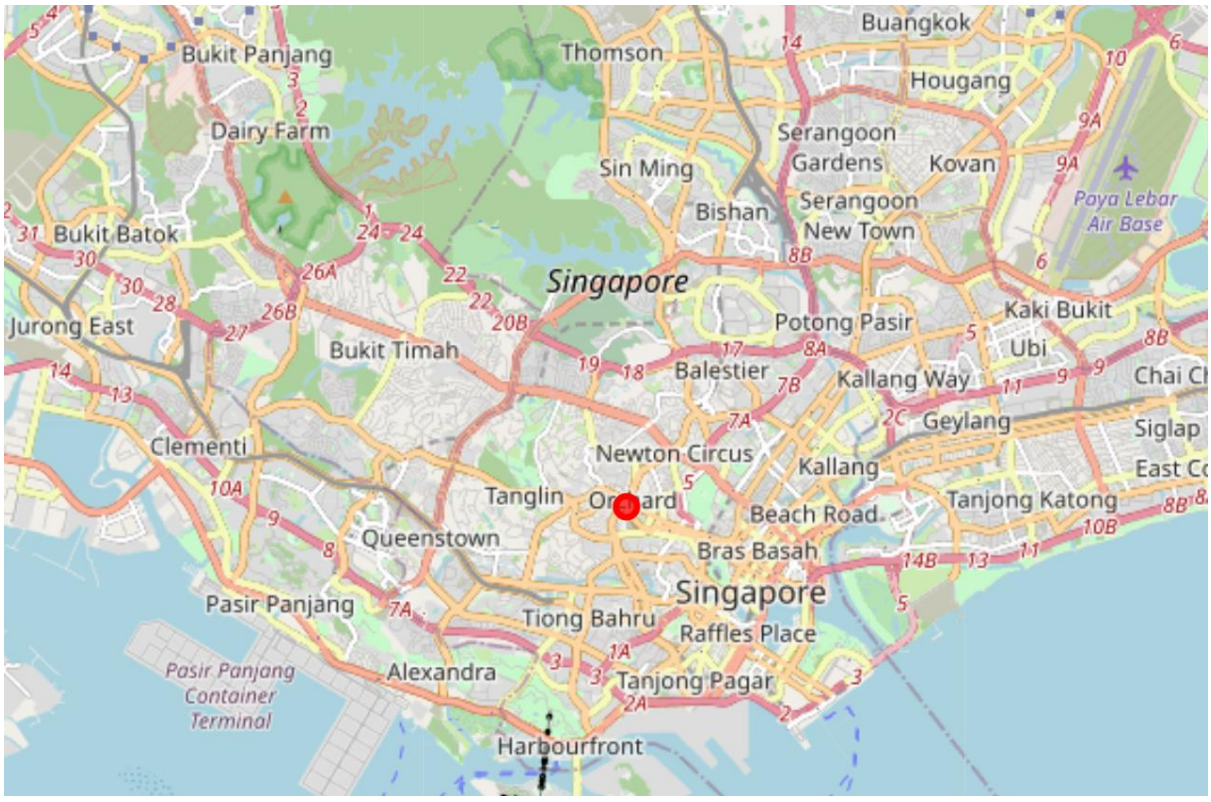

Figure 4: Clustering results for Orchard with the neighbourhoods of Toronto

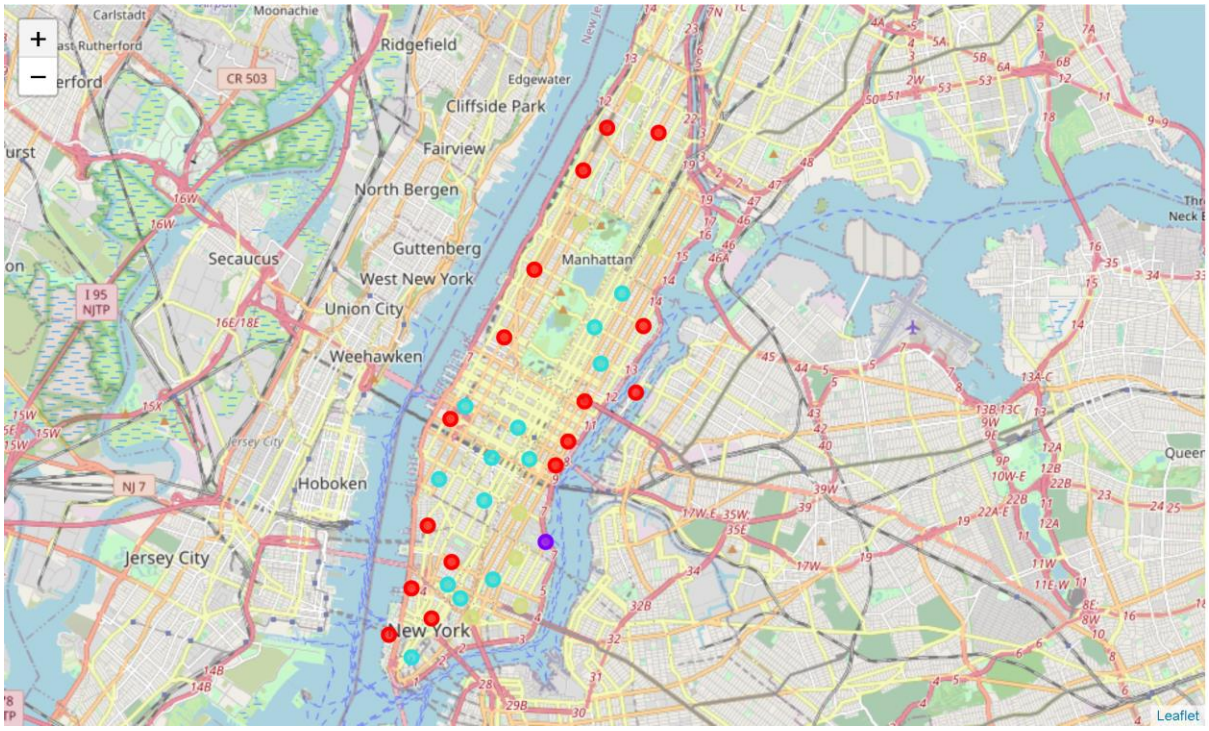

Figure 5: Clustering results for the neighbourhoods of Manhattan

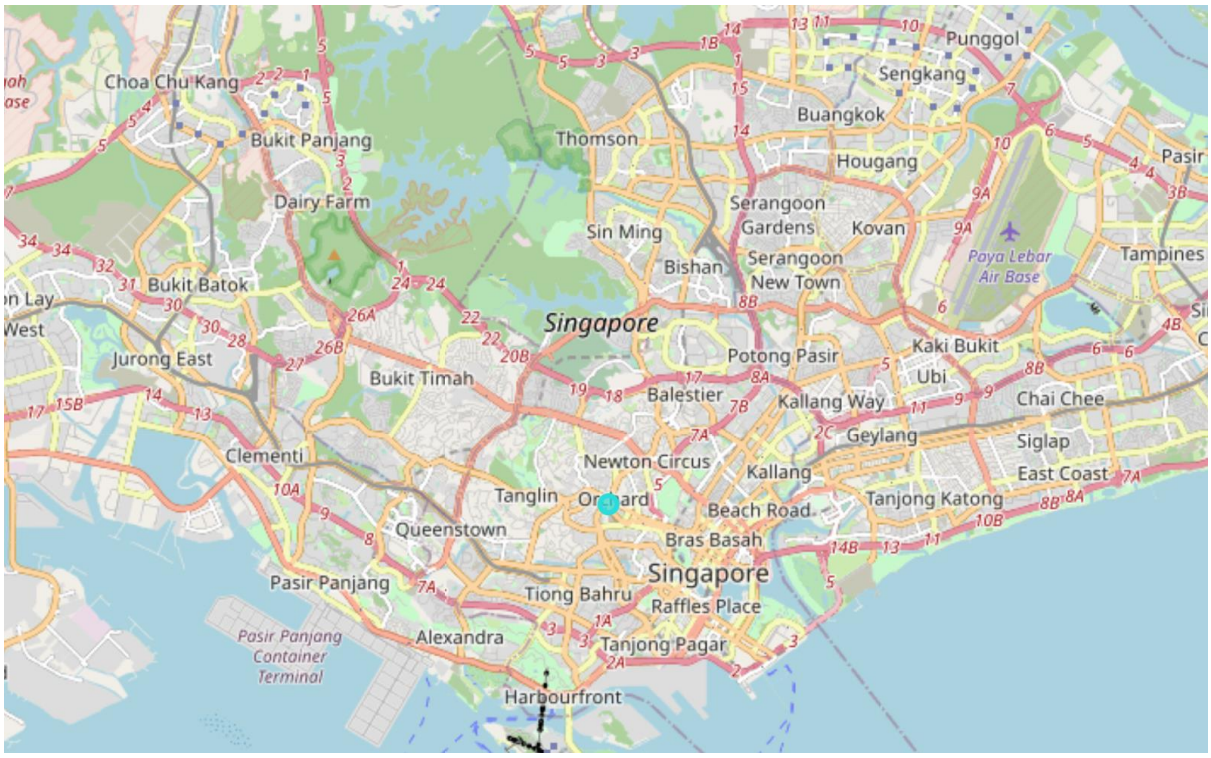

Figure 6: Clustering results for Ochard with the neighbourhoods of Manhattan

The results from the clustering of both the Toronto and Manhattan datasets are shown above in figures 3 – 6. Overall, the Orchard store was clustered with 14 other neighbourhoods in Manhattan and was clustered with 33 other neighbourhoods in Toronto. From this, we can deduce that there would be more

opportunities for Example pte ltd for new store locations if it expands to Toronto instead of Manhattan.

	Borough	Cluster Labels	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
0	Singapore	0	Boutique	Sushi Restaurant	Bakery	Hotel	Cosmetics Shop	Shopping Mall	Japanese Restaurant	Asian Restaurant	Bubble Tea Shop	Chinese Restaurant
0	Downtown Toronto	0	Coffee Shop	Bakery	Pub	Park	Breakfast Spot	Restaurant	Café	Theater	Dessert Shop	Chocolate Shop
1	Downtown Toronto	0	Coffee Shop	Sushi Restaurant	Yoga Studio	College Cafeteria	Beer Bar	Smoothie Shop	Sandwich Place	Burger Joint	Burrito Place	Café
2	Downtown Toronto	0	Clothing Store	Coffee Shop	Japanese Restaurant	Cosmetics Shop	Bubble Tea Shop	Italian Restaurant	Middle Eastern Restaurant	Café	Fast Food Restaurant	Bookstore
3	Downtown Toronto	0	Café	Coffee Shop	Restaurant	Cocktail Bar	Gastropub	American Restaurant	Clothing Store	Italian Restaurant	Moroccan Restaurant	Cosmetics Shop
4	East Toronto	0	Trail	Pub	Health Food Store	Dog Run	Department Store	Dessert Shop	Diner	Discount Store	Distribution Center	Women's Store
5	Downtown Toronto	0	Coffee Shop	Cocktail Bar	Beer Bar	Seafood Restaurant	Restaurant	Pub	Bakery	Café	Cheese Shop	Grocery Store
6	Downtown Toronto	0	Coffee Shop	Café	Italian Restaurant	Sandwich Place	Japanese Restaurant	Salad Place	Department Store	Restaurant	Bubble Tea Shop	Burger Joint
7	Downtown Toronto	0	Grocery Store	Café	Park	Italian Restaurant	Baby Store	Candy Store	Athletics & Sports	Diner	Restaurant	Nightclub
8	Downtown Toronto	0	Coffee Shop	Café	Restaurant	Thai Restaurant	Hotel	Deli / Bodega	Gym	Clothing Store	Pizza Place	Sushi Restaurant
9	West Toronto	0	Bakery	Pharmacy	Grocery Store	Middle Eastern Restaurant	Music Venue	Park	Pet Store	Pizza Place	Café	Brewery
10	Downtown Toronto	0	Coffee Shop	Aquarium	Hotel	Café	Fried Chicken Joint	Restaurant	Italian Restaurant	Sporting Goods Shop	Brewery	Scenic Lookout
11	West Toronto	0	Bar	Asian Restaurant	Men's Store	Restaurant	Café	Vegetarian / Vegan Restaurant	Coffee Shop	Yoga Studio	Record Shop	Pizza Place
12	East Toronto	0	Greek Restaurant	Italian Restaurant	Coffee Shop	Restaurant	Ice Cream Shop	Furniture / Home Store	Fruit & Vegetable Store	Pub	Pizza Place	Lounge
13	Downtown Toronto	0	Coffee Shop	Café	Hotel	Restaurant	Seafood Restaurant	Japanese Restaurant	Italian Restaurant	Salad Place	American Restaurant	Beer Bar
14	West Toronto	0	Café	Breakfast Spot	Coffee Shop	Grocery Store	Furniture / Home Store	Bar	Stadium	Bakery	Italian Restaurant	Climbing Gym
15	East Toronto	0	Sandwich Place	Fast Food Restaurant	Pizza Place	Pet Store	Pub	Italian Restaurant	Food & Drink Shop	Restaurant	Ice Cream Shop	Fish & Chips Shop
16	Downtown Toronto	0	Coffee Shop	Café	Restaurant	Hotel	American Restaurant	Gym	Italian Restaurant	Japanese Restaurant	Deli / Bodega	Seafood Restaurant
17	East Toronto	0	Café	Coffee Shop	Bakery	Gastropub	American Restaurant	Brewery	Stationery Store	Ice Cream Shop	Fish Market	Italian Restaurant
22	West Toronto	0	Park	Mexican Restaurant	Café	Thai Restaurant	Grocery Store	Fried Chicken Joint	Convenience Store	Music Venue	Diner	Cajun / Creole Restaurant
23	Central Toronto	0	Coffee Shop	Clothing Store	Yoga Studio	Bagel Shop	Gym / Fitness Center	Fast Food Restaurant	Diner	Mexican Restaurant	Pet Store	Chinese Restaurant
24	Central Toronto	0	Sandwich Place	Café	Coffee Shop	Park	History Museum	Liquor Store	Burger Joint	Indian Restaurant	Middle Eastern Restaurant	Flower Shop
25	West Toronto	0	Gift Shop	Breakfast Spot	Dog Run	Cuban Restaurant	Restaurant	Bookstore	Bar	Movie Theater	Italian Restaurant	Dessert Shop
26	Central Toronto	0	Sandwich Place	Dessert Shop	Gym	Italian Restaurant	Café	Pizza Place	Sushi Restaurant	Coffee Shop	Diner	Deli / Bodega
27	Downtown Toronto	0	Café	Restaurant	Bar	Italian Restaurant	Japanese Restaurant	Bookstore	Bakery	Coffee Shop	Chinese Restaurant	Dessert Shop
28	West Toronto	0	Café	Coffee Shop	Pub	Sushi Restaurant	Italian Restaurant	Tea Room	Grocery Store	Gourmet Shop	Sandwich Place	Restaurant
30	Downtown Toronto	0	Café	Coffee Shop	Mexican Restaurant	Vietnamese Restaurant	Bakery	Vegetarian / Vegan Restaurant	Bar	Dessert Shop	Grocery Store	Gaming Cafe
31	Central Toronto	0	Coffee Shop	Pub	Bagel Shop	Light Rail Station	Vietnamese Restaurant	Supermarket	Liquor Store	Sushi Restaurant	American Restaurant	Bank
32	Downtown Toronto	0	Airport Lounge	Airport Service	Airport Terminal	Boat or Ferry	Sculpture Garden	Airport	Airport Food Court	Boutique	Harbor / Marina	Bar
34	Downtown Toronto	0	Coffee Shop	Café	Italian Restaurant	Japanese Restaurant	Restaurant	Seafood Restaurant	Cocktail Bar	Beer Bar	Pub	Breakfast Spot
35	Downtown Toronto	0	Coffee Shop	Bakery	Restaurant	Café	Chinese Restaurant	Italian Restaurant	Pub	Pizza Place	Snack Place	Pet Store
36	Downtown Toronto	0	Coffee Shop	Café	Gym	Japanese Restaurant	Restaurant	Hotel	Steakhouse	Asian Restaurant	Salad Place	Seafood Restaurant
37	Downtown Toronto	0	Coffee Shop	Sushi Restaurant	Japanese Restaurant	Restaurant	Gay Bar	Café	Pub	Men's Store	Mediterranean Restaurant	Hotel
38	East Toronto	0	Light Rail Station	Yoga Studio	Pizza Place	Smoke Shop	Brewery	Farmers Market	Fast Food Restaurant	Burrito Place	Restaurant	Auto Workshop

Table 3: Clustering results for the Toronto Neighbourhoods

	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
0	Orchard	Boutique	Hotel	Sushi Restaurant	Bakery	Shopping Mall	Cosmetics Shop	Chinese Restaurant	Bubble Tea Shop	Indonesian Restaurant	Asian Restaurant
0	Marble Hill	Sandwich Place	Gym	Coffee Shop	Yoga Studio	Deli / Bodega	Supplement Shop	Steakhouse	Shopping Mall	Seafood Restaurant	Pizza Place
8	Upper East Side	Italian Restaurant	Bakery	Coffee Shop	Gym / Fitness Center	Yoga Studio	Wine Shop	Juice Bar	French Restaurant	Exhibit	Spa
10	Lenox Hill	Coffee Shop	Pizza Place	Italian Restaurant	Café	Sushi Restaurant	Cocktail Bar	Burger Joint	Gym / Fitness Center	Gym	Sporting Goods Shop
14	Clinton	Theater	Italian Restaurant	Coffee Shop	Gym / Fitness Center	Spa	Gym	Hotel	Thai Restaurant	Cocktail Bar	Sandwich Place
15	Midtown	Coffee Shop	Hotel	Theater	Bakery	Pizza Place	Japanese Restaurant	Clothing Store	Cuban Restaurant	Sushi Restaurant	Salon / Barbershop
16	Murray Hill	Hotel	Sandwich Place	Coffee Shop	Pizza Place	Gym / Fitness Center	Burger Joint	Steakhouse	Juice Bar	Bar	Grocery Store
17	Chelsea	Art Gallery	Coffee Shop	Ice Cream Shop	Café	American Restaurant	Bakery	Park	Cycle Studio	Cupcake Shop	Market
22	Little Italy	Pizza Place	Spa	Bubble Tea Shop	Italian Restaurant	Ice Cream Shop	Mediterranean Restaurant	Bakery	Thai Restaurant	Sandwich Place	Chinese Restaurant
23	Soho	Italian Restaurant	Mediterranean Restaurant	Sandwich Place	Coffee Shop	Ice Cream Shop	Bakery	Dessert Shop	Sushi Restaurant	Spa	French Restaurant
29	Financial District	Coffee Shop	Pizza Place	American Restaurant	Hotel	Café	Falafel Restaurant	Gym	Cocktail Bar	Sandwich Place	Event Space
30	Carnegie Hill	Coffee Shop	Café	Pizza Place	Yoga Studio	Wine Shop	Japanese Restaurant	Gym	Bookstore	French Restaurant	Shipping Store
31	Noho	Pizza Place	Coffee Shop	Italian Restaurant	Japanese Restaurant	Sandwich Place	Grocery Store	Mexican Restaurant	Wine Shop	Rock Club	Sushi Restaurant
33	Midtown South	Korean Restaurant	Hotel	Japanese Restaurant	American Restaurant	Dessert Shop	Coffee Shop	Burger Joint	Hotel Bar	Gym / Fitness Center	Café
38	Flatiron	Gym / Fitness Center	Café	Italian Restaurant	Yoga Studio	Park	Gym	Japanese Restaurant	Coffee Shop	Spa	Vegetarian / Vegan Restaurant

Table 4: Clustering results for Manhattan Neighbourhoods

5. Conclusion

In conclusion, the classification model suggests that Example pte ltd will have more success in its expansion into the North American region if it decides to set up stores in Toronto, Canada instead of Manhattan, New York. However, it should be noted that the Manhattan also remains a promising venue for expansion should Example ptd ltd decide on entering the United States market.