Confidential Customized for **Example Pte Ltd** Version 10

# North America Expansion



#### Introduction to the Problem

Example ptd Ite is looking to expand its business and stores into the North American region. Initial research has narrowed down the options to two possible cities: Toronto and Manhattan. However the strategic management is unable to reach a conclusion on which city to go from there and has contacted the Data analytics team for help. Analysis from previous expansions have shown that stores most similar to the Orchard store in Singapore yield the most success, so we will be comparing the different neighborhoods and clustering them to see which country has the most neighborhoods that are similar.

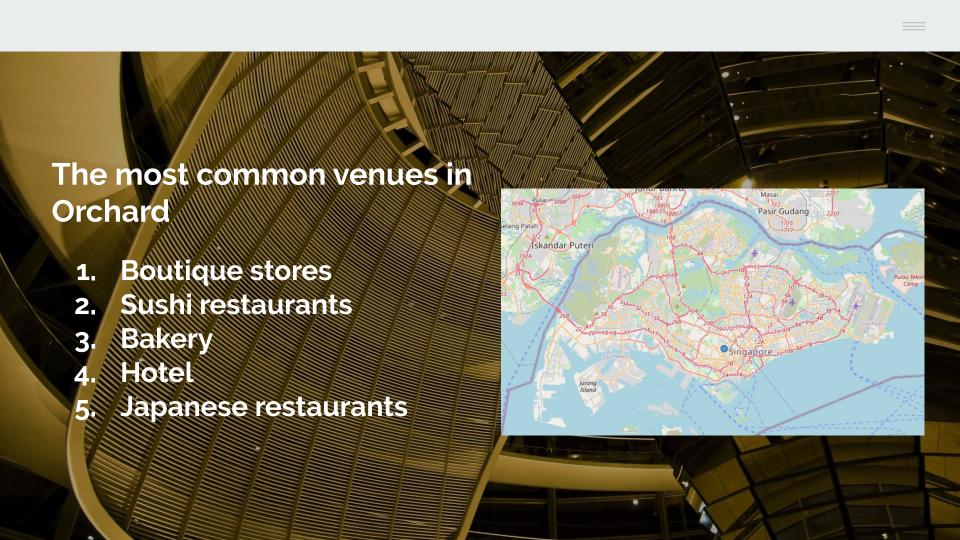


## **Data Acquisition**

The postal codes for the Toronto neighborhoods were first obtained by scraping from the wikipedia page.

Using, the postal codes, the geospatial data could then be obtained for the Toronto neighborhoods. The geospatial data for the Manhattan neighborhoods were already provided.

- The geospatial data, which contained latitudes and longitudes, were then used to obtain information about the venues within 500 meters of the respective neighborhoods
- These data were then grouped and encoded to obtain a table containing the relative frequencies of the various venue categories for each neighborhood



**Classification Models** 

K-means clustering was used to group the different neighborhoods based on the frequencies of venues.

### Results

#### Manhattan vs Toronto

Overall, the orchard store was clustered with 33 other neighbourhoods in the Toronto dataset and was clustered with 14 other neighbourhoods in the Manhattan dataset



There are more stores that are similar to the Orchard store in Toronto than in New York so Example pte ltd would have a better chance of expansion in Toronto than in New York.



# Thank you.

