REPORT – APPLIED CAPSTONE PROJECT

1. Introduction

This report is to help the companies and entrepreneurs that planning to expand their business, open new shops or franchise in Toronto, Canada. However, the main problem that majority of the companies and entrepreneurs facing is where is the best location to open a new shops or franchise and what is best shops or franchise to open in certain locations of Toronto, Canada. Moreover, companies and entrepreneurs also very concern about the number of competitors around there. In order to solve this problem, the data about the shops of each neighbourhood will be collected to show what are the current shops at each locations, how many similar shops, demand of the shops, what type of shops and more. The target audience of this report will be the companies, entrepreneurs or people that interested to expand their business, open new shops or franchise in Toronto, Canada.

1. Data

In order to solve the problem, the data that consisted of postal codes, neighbourhood and borough of Canada will be collected. The number of shops of each neighbourhood will be collected through Foursquare API by combining with the data that consisted of postal codes, neighbourhood and borough of Canada. There are 103 observations in the data. After combining the data, the data will be showed what are the shops at each neighbourhood and what type of shops are high demand. For example, location A having cafe A, korean restaurant, gym, cafe B, cafe C, Japanese restaurant, gym and more. As we can see from the list, cafe is the most popular in the location A, which is high demand but also many competitors. The end result will be showed in the list or table format, so that it is easy for people to understand and see it.

Data that consisted of postal codes, neighbourhood and borough of Canada : <https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M>

1. Methodology

In order to make the final result more accurate, some exploratory data analysis will be processing such as cleaning and modification.

* 1. Exploratory Data Analysis
     1. Data Cleaning

Firstly, there are many missing values under the borough variable. In order to produce the best result, the cells that having “Not assigned” under borough will be ignored and process with the cells that have an assigned borough. On the other hand, there are some missing value can be noticed in the neighbourhood variable. The neighbourhoods that having missing values will be assigned the same borough to the neighbourhood for that column. For examples, the value of the Borough and the neighbourhood columns will be Queen's Park.

* + 1. Data Modification

There are many neighbourhood that having same postal code area which means that more than one neighborhood can exist in one postal code area. For instances, postal code area M5A has two neighborhoods which is Harbourfront and Regent Park. The neighbourhoods that having the same postal code area will be combined together with a comma and put in the same postal code area. Moreover, in order to utilize the Foursquare location data, two new variables which is latitude and the longitude coordinates of each neighbourhood need to add into the cleaned data.

* 1. Machine Learning

The machine learning method that will be using is K means clustering. This is because this method able to cluster the result into group which is like what are the shops at that location. By looking at the cluster, the user can understand it easily and make the best decision.

1. Result and Discussion

The final result is in the table format which listed out all the top 10 most common shops in that location by the postal code area, borough, neighbourhood, longitude and latitude. The user able to understand the table easily by looking at the location they want. The final result may not be accurate in the future as there are many new shops, shops closed down, changing of demand and more factors will affect the locations. Hence, some update needed to do every year in order to help the companies and entrepreneurs for their business.

1. Conclusion

In conclusion, this report able to help the companies and entrepreneurs that planning to expand their business, open new shops or franchise in Toronto, Canada. This will increase the job opportunity and attract many customers to visit.