**COFFEE SHOP**

**RECOMMENDER SYSTEM**

# - IBM Applied Data Science Capstone Project

Contents:

## 1. Introduction 2. Data 3. Methodology 4. Result 5. Recommendation 6. Conclusion

**INTRODUCTION :**

**PROBLEM:**

“to develop a recommender system that will help the restaurant manager to find the best suitable place to open a Coffee Shop. “

**PROBLEM BACKGROUND:**

Toronto is the provincial capital of Ontario and the most populous city in Canada, with estimated population of 2,956,024 and an estimated population of 6,341,935 in the Toronto region. Toronto is an International centre of business, finance, arts, and culture. Its large population of immigrants from around the globe has also made Toronto one of the most multicultural and cosmopolitan cities in the world.

The city has many restaurants, coffee shops, cafe, hotels. The variety of food items are provided by these shops. One of the popular items is a Coffee. Thus, there are many Coffee Shops in the various areas of the city. Therefore, if someone decides to open a Coffee Shop in the city, he would select the best suitable place for the shop. To get the information about this suitable place, the recommender system can be used. This recommender system will get the data, analyse and visualize it; and then provide the best place to open a shop or restaurant.

**PROBLEM DESCRIPTION:**

If the manager decides to open a Coffee Shop in Toronto, then following questions must be addressed –

1. Which place is the most suitable and popular for the Coffee Shop?
2. What type of Coffee should be provided? What type is preferred by people in that area?
3. What type of people live in that area (students, company employees, etc)?
4. How many similar shops are present in that area?
5. What other specialities should be provided to attract customers?
6. What should be the cost of the Coffee provided?

To get answers of above questions, a recommender system can be implemented. This system can help the managers to decide the best place to open a Coffee Shop based on the rankings of places.

The recommender system should be able to answer the all questions and provide best solution. It should provide – popular items in that area, target audience, their average income, cost estimation, etc.

**TARGET AUDIENCE:**

Target audience for this system are the managers or people who want to open a Coffee Shop in the specific city or area. These people expect the place which is most popular and well known in the city. They also need the information about popular items, categories, other specialities, etc. Thus, using this recommendation system, the managers can decide the most suitable place for the Coffee Shop.

**DATA:**

To build a recommender system to find the best suitable place for a Coffee Shop, huge amount of data is required. The datasets are used to analyse the data, visualize the data, and the finally to get suitable solution to the problem.

To open a shop, following things are required –

1. Geographical coordinates of the area
2. The population of the neighbourhood
3. The type of people in the neighbourhood
4. Average income of the people nearby that area
5. The preference of people towards the type of food
6. Other service details such as juice, transport, taxi, etc.

The more popular the area, more chances to successfully provide service to people. The type of people (students, or employee) plays the important role to decide what type food should be provided.

The average income helps to decide the cost of the food items. The preference of the people should be taken in consideration. Other special services such as juice help to increase the business opportunities.

Thus, to get the solutions, a large amount of data is required.

**DATA COLLECTION:**

The names, postal codes and addresses of the places around the city can be obtained from various websites such as Wikipedia. The geographical coordinates data can be obtained from the various open-source websites such as Wikipedia, Google Maps, Government websites, census report websites, etc.

The population of the area can be easily obtained by searching on the websites.

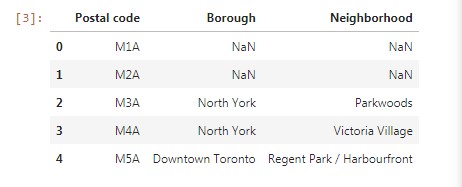
The average income data can be obtained from Wikipedia, or other sources.

Foursquare API can be used to get the information about the nearest locations.

This data can be used to visualize, clustering the results.

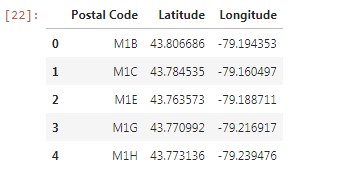
**METHODOLOGY**:

The first task was to get the data of neighbourhoods in Toronto. The list of neighbourhoods of Toronto was extracted from the Wikipedia page. The list consists of the postal code, borough and the area name.

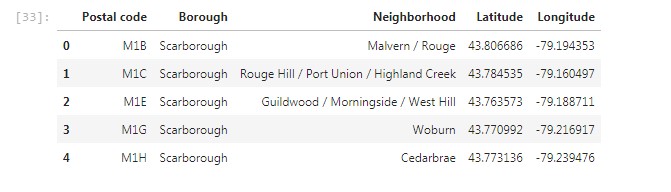


Then I used the pandas HTML table scraping method for web scraping.

Then to get the longitude and latitude coordinates of the areas, the CSV file was used.



After matching the areas and the coordinates, I visualized the map of Toronto using the Folium library package.



Then to get the list of top nearest venues, I used the Foursquare API, using my ID, secret key and version. Then I obtained the names, addresses, longitude, latitude of the nearest venues using it.

Then I grouped the venues by their categories. Then I selected the category as “Coffee Shop “. Then I used the K-Means clustering method to form the clusters of the data.

Then by analysing the results, the final results were obtained.

**RESULT**:

The clusters obtained using the K-Means clustering method:

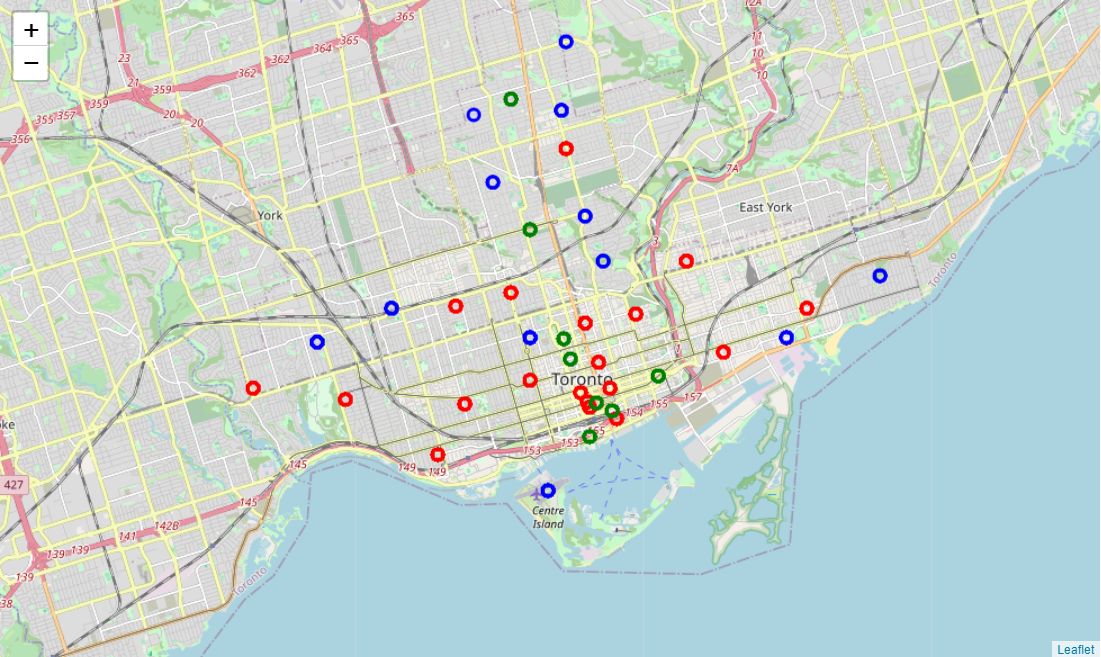


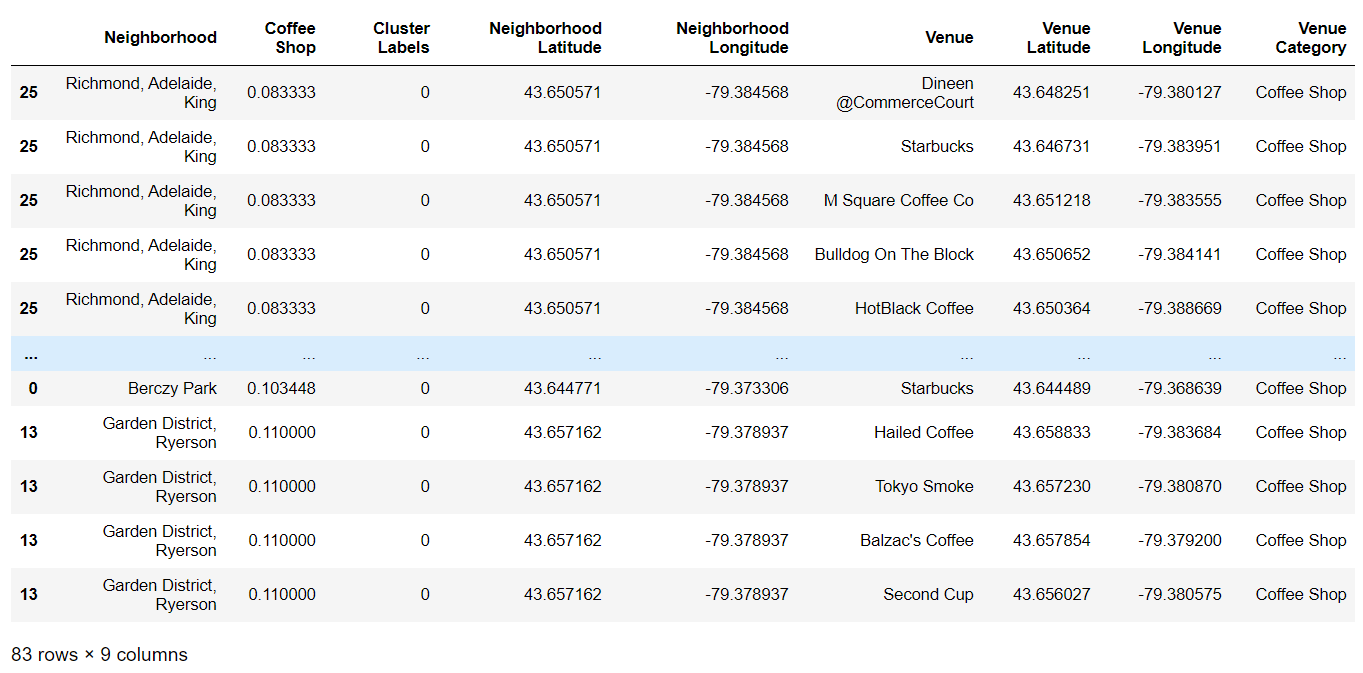
Table - Cluster 0: 

Table – Cluster 1:

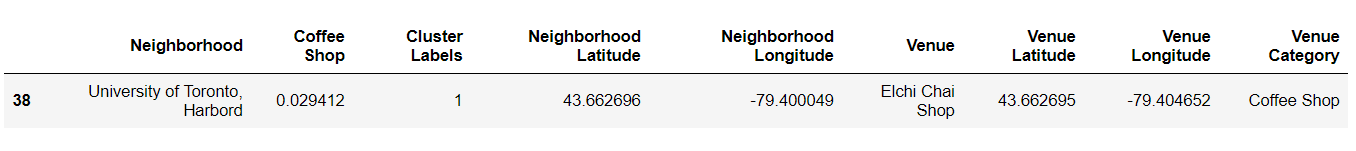


Table – Cluster 2:



**RESULT**:

By analysing nearby venues, we can conclude that the cluster 1 does not have many Coffee shops in that areas. Thus, it would be suitable to select these locations for opening Coffee shops.

**RECOMMENDATION**:

Therefore, locations like Rosedale, Roselawn , University of Toronto Harbord etc will be good to open a new Coffee Shop.

**CONCLUSION**:

The recommender system correctly recommends the most suitable place to open a Coffee Shop. Thus, it can provide good results to the users of the system. The system can also be used as recommendation system for opening the restaurants, coffee shops, street food shop, etc. Using this method, the recommendation system for malls, theatres, shops can also be designed.