Find a suitable location to open a restaurant in Toronto

**Introduction:** As the population is constantly increasing, the requirement of Different services are also increasing. Hence, Opening a Restaurant or a chain of Restaurant is a very common business idea For any entrepreneur nowadays. The Success and profit of the restaurant are dependent on several factors like locality, type, services etc. To Maximize the profit of the Restaurant, It is important to find the answer of a few fundamental questions. One of the very first problems is to find a suitable locality for the new Restaurant as location plays a big role in the success of Restaurant. In this project, I am trying to solve this problem using Data Science and Machine learning Algorithm.

**Business Problem:** The objective of my project is to find a suitable location to open a new restaurant in Toronto, the financial capital of Canada. I have used data science methods and machine learning algorithms using the programming language python in this project. The goal of this project is to provide the answer to the business question "In Toronto if an entrepreneur wants to open a new restaurant, which neighbourhood/neighbourhoods will be suitable for this?

**Data:** We dont use to see ongoing and successful restaurant everywhere in a city. The concept behind this project is, As locality of a restaurant is a key factor if we can identify the similar locality based on different Characteristic of those localities, we can choose our suitable location as well.

As an example, If neighbourhood N1, N2, N3 is similar and if Restaurants in N1 and N2 are successful then there is a high possibility that N3 is also a suitable location for Restaurant.

In this project, we have used the below data to find the similar Neighbourhood.

i. Different neighbourhood information of Toronto City along with their popularity, income and other details: We have extracted the data from The official website for the City of Toronto (www.toronto.ca). please Note, up to the year 2016 data is available here hence we have used the latest data (i.e. the data of 2016) only.

ii. Latitude and Longitude of all neighbourhoods: we have extracted the Latitude and Longitude data of all neighbourhoods using Geocoder.

iii. Different venue details based on all neighbourhoods: we have extracted the data based on Foursquare API

**How this Data will resolve the problem**

Using this data I am going to create clusters of neighbourhoods based on its feature with the help of K-means algorithm. Let me explain what logic I have used to achieve our goal.

Let's say, we have 5 neighbourhoods such as N1, N2, N3, N4 and N5 and number of restaurant in those neighbourhoods are 10, 2, 2, 13, 4 respectively

And After using a clustering algorithm based on its features ( excluding the restaurant number), let's say we have 2 clusters such as ( N1, N3, N4 ) and (N2, N5 )

As N1, N3 and N4 are similar and we have a high number of successful restaurants in N1 and N4, we can expect N3 is also a good location for a restaurant with less competition. For this I have not used Restaurant number is normalization as I am not going to use it.

From our initial Dataset, I will get the details of neighbourhoods. Then I will extract geocode of neighbourhoods and venue details around the neighbourhood. At last, I will use machine learning Algorithm to form cluster.

Then I will analyze (the way I have describes above) clusters to find the suitable location