**Background**

A group of 4 friends have come together and started a small band. They are now looking to setup a small music cafe in a locale where people will be able to come and watch them perform or record music. The music café has to be in a place where people of all ages conglomerate. The genre of music the band plays appeals to both young and old. The band wants their music cafe to be in a locale where people gather for passing their time such as a park, or where people come to simply hangout with their friends such as cafes and coffee shops.

**Problem** : The music band called the Wonders, is not yet profitable as the group of friends are unable to gain the necessary audience. The Wonders want to setup a cafe and be able to attract crowd, while also leveraging a good location where the rent would not be excessive as the Wonders don’t have much money right now. They are looking for the perfect location in the city of Toronto in a borough in Canada by the name of ‘Etobicoke’

**Data Description:**

The neighborhoods data has been pulled from the Wikipedia page provided. All the data including the latitude and longitude was made available with the course material itself. The neighborhood data enabled in superimposing all the neighborhoods available in the borough of ‘Etobicoke’ in the city of Toronto. The foursquare API data helped in location venues.

**How it solves the problem**

The foursquare data gives a neighborhood wise location of key areas of interest in a city such as a park, hotel, airport, café, coffee shops, restaurants etc. As mentioned in the background section, the music band by the name of Wonders would like their cafe to be in an area where people hang out often and pass their time, the data leverage from foursquare will have this information.

**Exploratory data analysis and K-means algorithm used**

The borough Etobicoke was selected as the location where the music band called the Wonders wanted to open up a music café. In order to find the best suitable location, all neighbourhoods were explored by means of K-means algorithm. The algorithm enabled an easy search of neighbour hoods which have high level of commercial and business outlets. In this case Cluster 0

**Results**

To select a suitable location for the music café where a locale with diverse groups of people should be visiting was required. Using data from Foursquare API, Alderwood from ‘Cluster 0’ looked like the most suitable location given the summary which suggested the place as popular. Availability of a large number of hangouts and eating places meant a larger population gathers. Of the 24 venues displayed by Foursquare none were reported as Cafés. A good business opportunity presented itself whereby a café plus a new music band combine could attract large number of people.

**Discussion**

Cluster 3 has a baseball filed which also means it’s a large public gathering spot, conversely it has fewer number of eateries as compared to cluster 0. A good location thus for either a pizza place or foreign cuisine.

**Conclusion**

Cluster analysis reveals that a large number of neighborhoods are cluster into 1 single cluster. One can infer the likes and dislikes of the majority of the population in the city of Etobicoke from the K-means analysis. This can be narrowed down right to neighborhood as well by leveraging the data from Foursquare. Using similar such datasets, it is easy to infer the tastes of a certain section of the population, further any new startup can also leverage this type of data to start a new business and gain quick traction and revenue as well.