**Coursera Capstone - The Battles of Neighborhoods**

***Recommendation on the Toronto neighborhoods of choice for ABC Grocery***



**1.0 Introduction:**

Toronto is a well-developed capital city of Canada, with lots of business opportunities and business friendly environment, it has no issue in attracting many different players into the market.

However, that also means the market is highly competitive and as a well-developed city, the cost of doing business is also one of the highest in the country. And thus, any new business venture or expansion in the country needs to be reviewed carefully and strategically targeted so that the return on investment will be sustainably reasonable and more importantly the investment can be considerably less risker.

**1.1 Problem Description:**

This is clearly a problem that an online grocery retailer (i.e. ABC Online Grocery) needs to review and resolve as part of their new business venture in the country. As a startup though well-funded, they need to choose their first starting location carefully for the points highlighted above and more importantly, if this is successful, they will be able to replicate the same success fairly quickly; so, first mover advantage is critical for this business and thereby the choice of location (i.e. neighborhood) is important to them.

**1.2 Target Audience:**

To solve this problem, data scientist team led by myself has been engaged by ABC Online Grocery. The objective is to locate and recommend to the management which neighborhood will be the best choice to start off their first online grocery delivery services. The management also expects to understand the rationale of the recommendations in the report.

**1.3 Success Criteria:**

The success criteria of this project will be a good recommendation of the neighborhood choice in Toronto to the management of ABC Online Grocery based on 2 key factors; lack of grocery stores available (less competition) and higher number of residences presented (higher demand).

**2.0 Data Description:**

As we need to explore, segment, and cluster the neighborhoods in the city of Toronto, the Toronto neighborhoods data is key for this project. Unfortunately, the data is for the Toronto neighborhood data is not widely available on the Internet in the structured format, hence we need to scrap it through an existing Wikipedia page exists that has all the information we need to explore and cluster the neighborhoods in Toronto. The data should contain the coordinates for each of the neighborhood in Toronto that will help us to further obtaining more information critical for this project. We will also like to obtain the key information like below; such as number of residences information for each neighborhood which is one of the key factors for the neighborhood of choice in the final recommendation report. The data needs to be clean up and eventually in a structured format like the example below.

1. Neighborhood Name
2. Neighborhood Latitude
3. Neighborhood Longitude
4. Number of residences in each neighborhood.



**2.1 Data Features:**

We will be leveraging on features in a reliable location information provider such as the Foursquare.com to explore the various types of venues and its categories available in each neighborhood. We will also need to understand the trending of these venues in the respective neighborhood. The information obtained per neighborhood will be as such like below and has to be in a structured format:

1. Neighborhood
2. Neighborhood Latitude
3. Neighborhood Longitude
4. Venue Name
5. Venue Category
6. Venue Latitude
7. Venue Longitude



**3.0 Methodology**

*Methodology section which represents the main component of the report where you discuss and describe any exploratory data analysis that you did, any inferential statistical testing that you performed, and what machine learnings were used and why.*

We will also use clustering techniques such as the K-Clustering to segment and cluster these neighborhoods so that we can group them together to understand their similarities and what best we can do for these types of neighborhoods.

With all these features, techniques and data, we will then be able to come up with a best recommendation to the management of ABC Online Grocery to their problem which is where is the best neighborhood for them to first start off to offer their services. For an example, we will not want to enter a neighborhood whereby there is already a high concentration of grocery stores available or there is a high trending of such stores upcoming in the neighborhood; we will like to recommend a neighborhood where we know that there will be a higher demand of such delivery service due to the lack of supply in that area.

**4.0 Result**

*Results section where you discuss the results.*

**5.0 Discussion**

*Discussion section where you discuss any observations you noted and any recommendations you can make based on the results.*

**6.0 Conclusion**

*Conclusion section where you conclude the report.*