**Applied Data Science
Capstone project report
(IBM, Coursera)**

The Battle of Neighbourhoods

City of Toronto

Navpreet Sharma

Contents

1. Introduction

There are various ways of expanding a business, and companies can choose different growth strategies to reach their goal. But one the most common strategy is targeting new customers in new locations for the existing products. The following project takes into consideration the neighbourhoods of a city, and finds the similarity between them in order to develop insights for businesses to choose the right new market for their product.

1.1. Business Problem

To find a list of most apt neighbourhoods in the city of Toronto to open a new branch of a successful Chinese restaurant running in Maple Leaf. This would be done by finding neighbourhoods that are most similar to Maple Leaf with respect to the venues distribution across these neighbourhoods and then using the demographic and income data of each neighbourhood to get the top ten list.

2. Data

In order to the analysis on the city of Toronto, the following data is needed:

- **Geospatial data** – this consists of neighbourhood names and their geographical coordinates.

  Source - https://open.toronto.ca/dataset/neighbourhoods/

- **Venues data** – this consists of a list of all the venues, such as parks and restaurants, present in each neighbourhood and information related to these venues, such as category, location, reviews and tips. This data will be sourced from the Foursquare API by making a query to it (using 'explore' endpoint on each neighbourhood) with our credentials. This data is in the form of a JSON object and the 'venue category' part of each venue is used to build the data for analysis.

- **Demographic data** - this consists of total population and the population of people with Chinese origin in each neighbourhood.

  Source - https://open.toronto.ca/dataset/wellbeing-toronto-demographics/

- **Income data** - this consists of median household income of each neighbourhood.

  Source - https://open.toronto.ca/dataset/wellbeing-toronto-demographics-nhs-indicators/