IBM Applied Data Science Capstone – Battle of the Neighborhoods

INTRODUCTION

Toronto and Vancouver represent the first and third largest Census Metropolitan Area's (CMA's) in Canada with a 2016 census population of 5,928,040 and 2,463,431 respectively. Toronto is located in eastern Canada on the shores of Lake Ontario and Vancouver is located in Western Canada on the shores of the Pacific Ocean (Figure 1).

The Toronto CMA has the largest proportion of foreign-born residents (46%) as a share of the total population out of all metropolitan areas with the largest groups being Asian, African, Latin/Central/South American. In contrast the Vancouver CMA showed that 48.6% of the population was of European heritage with the largest visible minorities being Asian, and Latin Americans.
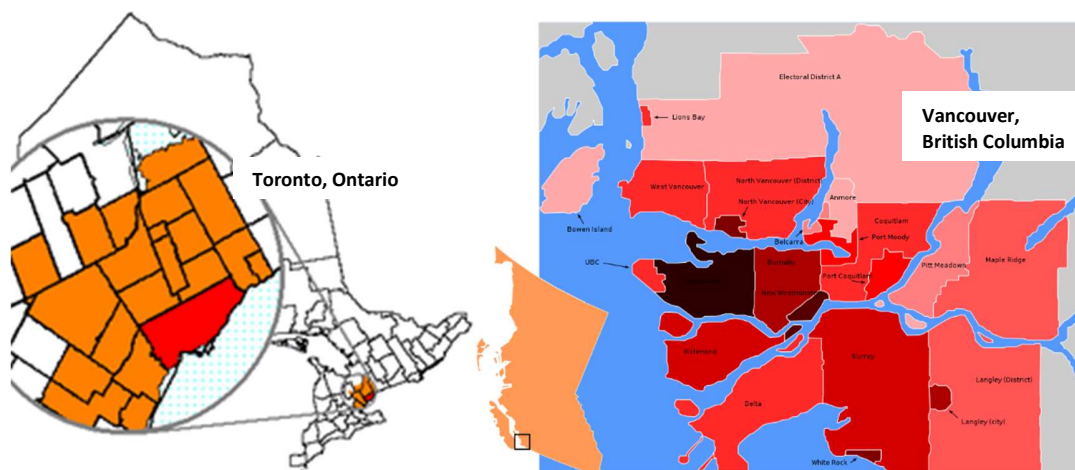

Figure 1. Maps of Toronto and Vancouver Census Metropolitan Area's.

PURPOSE

The purpose of this project is to fulfill the capstone requirements of the IBM Applied Data Science Certificate by applying the data wrangling, visualization, and analysis (K-means clustering) methods. This project will examine the similarity and dissimilarity of the CMA's of Toronto, Ontario and Vancouver, British Columbia Canada.

METHODS

Using Forward Sortation Area's (FSA's) postal codes to identify the municipalities/neighborhoods geographic location of both Toronto and Vancouver CMA's, Foursquare data for each CMA will be queried and analyzed by summarizing the venue characteristics of the municipalities/neighborhoods and applying the k-means clustering algorithm. The venue characteristics of the Toronto and Vancouver municipalities/neighborhoods will then be compared.

DATA

The following data will be used:

- List of FSA's (Postal Codes) and neighborhoods in the CMA's of Toronto, Ontario and Vancouver, British Columbia Canada.
    - Toronto, Ontario:
        - https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
    - Vancouver, British Columbia:
        - https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_V
- Latitude and Longitude of the postal codes
    - Toronto, Ontario:
        - http://cocl.us/Geospatial_data
    - Vancouver, British Columbia:
        - https://www12.statcan.gc.ca/census-recensement/2011/geo/bound-limit/bound-limit-2016-eng.cfm
- Foursquare venue data of Toronto and Vancouver queried by postal code.
        - https://foursquare.com/developers/login?continue=%2Fdevelopers%2Fapps