IBM Data Science Capstone Final Assignment

Week 4

Neighborhood Similarities Between Cities

Introduction/Business Problem and Data Description

Contents

Introduction

At some point every Business Owners think about expansion, once their business is locally consolidated the next step is move into international waters. This decision comes with a new set of challenges and opportunities but also with a higher risk since this is literally a New Territory. An analysis of similarities between neighborhoods will allow comparing known successful locations against new candidates, and make and inform decision to where open a new business branch.

Business Problem

Our client is a successful owner of several boutique French Cafes in New York City, which is currently evaluating the possibility to start a new venture in Toronto. For that, we need to answer the following question: which neighborhoods in Toronto have the biggest similarities with New York neighborhoods were the customer enterprise has proved itself successful?.

The customer successful venues are mostly located in the neighborhoods of Washington Heights and Manhattan Valley.

Target Audience

The current project will be presented to the Business Owner and Directive Board, it would require a high level presentation of the analysis in order for the executives to appreciate the validity of the final recommendations.

Additionally the results can be of interest to any business owner that want to expand their business into Toronto city.

Data

For this project, the following data will be used:

- New York City neighborhood data.
- Toronto neighborhood data
- Venue Data for FourSquare, that will allow us to cluster and classify neighborhoods together and identify the similarities between both cities.

Sources and Extraction

Data 1:

New York City neighborhood data will be extracted from a dataset available for free on the web:

https://geo.nyu.edu/catalog/nyu_2451_34572

This information is structured in 5 boroughs and 306 neighborhoods and contains the latitude and longitude coordinates of each neighborhood

Data 2:

Toronto City information will be extracted from the city's Wikipedia page:

https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M,

We'll scrap the data using BeutifulSoup, and we were leveraging the Google Maps Geocoding API to get the latitude and the longitude coordinates of each neighborhood.