IBM Professional Certificate in Data Science Capstone

Relocation Recommendation system based on Neighborhood similarity

By:

Mfongeh Noel

April 10, 2020

Introduction

This project is aims to aid the user with choosing an ideal location for relocating an office space, store, home to another city with similar characteristics. It uses neighborhood segmentation and clustering to group neighborhoods in the destination city and on this basis, recommend an array of similar or better neighborhoods to choose from.

The similarity between neighborhoods is measured by comparing the venues in the neighborhoods. Therefore, two neighborhoods with the same type of venues would be considered similar. The type of venues accommodated in this study include places to shop, eat, drink and places for sight seeing

This study uses foursquare API and many other data sources to acquire location data. These are described in greater detail in the next section. For this study, we compare the neighborhoods of Toronto and Manhattan but the same concept can be extended to any pair of cities in the world.

Target Audience. The specific target users of this kind of application would include small start ups and employees who wish to relocate for better professional opportunities without compromising on the cost or quality of life

2 Data

2.1 Foursquare Places API

Foursquare is a social location service that allows users to explore the world around them. Users can download the Foursquare application to their iPhone, Blackberry, or Android phone and sign up for free, then connect their Foursquare accounts to their other social media accounts.

The Foursquare API allows application developers to interact with the Foursquare platform. The API itself is a RESTful set of addresses to which you can send requests and ger responses. The API allows querying places and users, exploring popular places, and checking out reviews and photographs for these places.

In this project we will be using Foursquare Places API to identify the most popular venues for each neighborhood. This is possible by the use of a explore call that returns a list of recommended locations in a specified area.

2.2 Other Data Sources

New York Neighborhoods Data was obtained from NYU Spatial Data Repository. The repository was titled 2014 New York City Neighborhood Names.

Toronto Neighborhoods Data was scraped from a Wikipedia page titled List of Postal Codes of Canada : M.