Peer-graded Assignment:

Capstone Project - The Battle of Neighbourhoods

Introduction/Business Problems

Fabio Melle

Background

Scarborough (Ontario), is a municipality in Canada and a suburb of Toronto. Located above the Scarborough Bluffs, occupying the eastern part of the city. Scarborough is bordered by Victoria Park Avenue to the west, Steeles Avenue to the north, Rouge River and Pickering City to the east and Lake Ontario to the south. The city is named after Scarborough (United Kingdom).

The city is one of the most multicultural of the Greater Toronto Area, hosting various religions and cultures. The city has also been declared the greenest of the Greater Toronto Area.

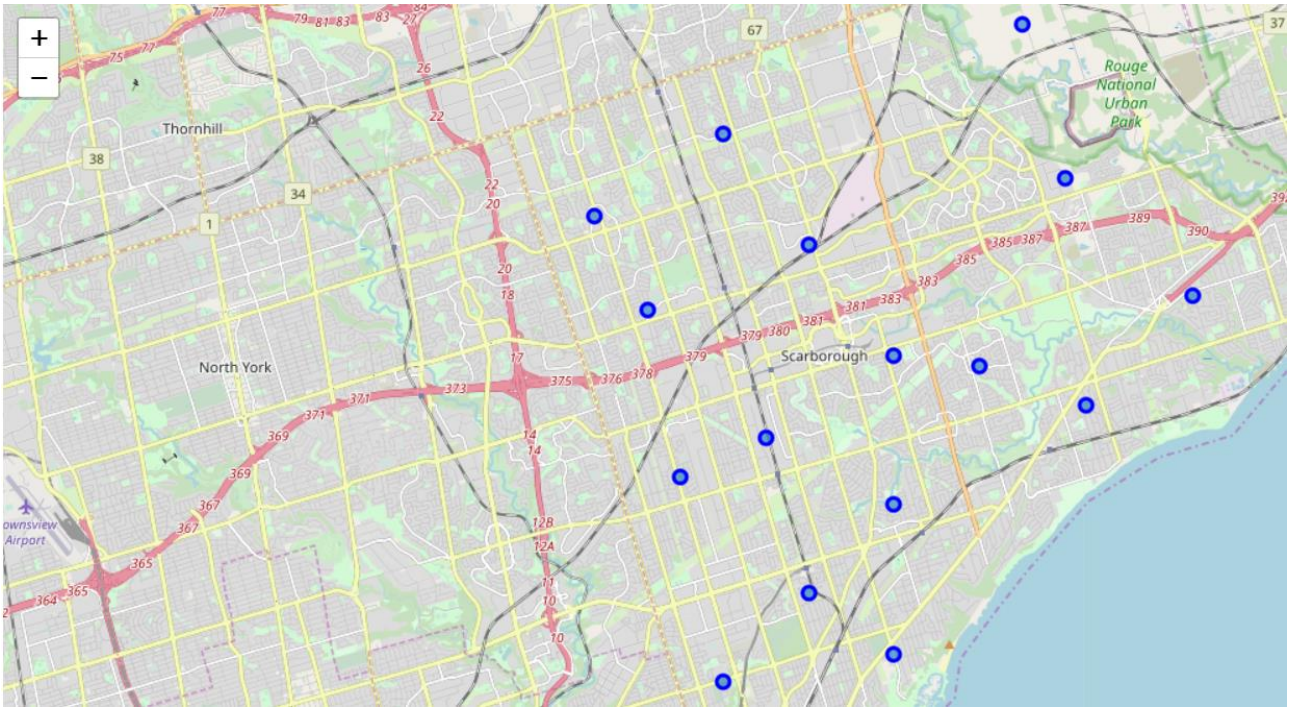

Figura 1 - Scarborough and its Neighborhoods

Description

Some Italian people have realized that new job opportunities have opened up in Scarborough, having relatives living there. So they decided to move to this city. Before leaving, they would like to know more about this city and its neighborhoods so that they can evaluate where to live. Before leaving, they would like to have more information about this borough and its neighborhoods in order to evaluate where to go to live.

Goal

The aim of this work is to give these families the opportunity to settle in a good neighbourhood by taking into account some evaluation parameters such as the services that are inside, the quality of education guaranteed by schools and the average cost of housing.

Foursquare API:

As the main source of data collection I used the Foursquare API as it has a database with a capacity of millions of places, especially the place API which offers the ability to perform location searches, location sharing and detailed information about a particular company.

By using the Foursquare Developer API with my account I will be able to get places close to the districts in the district. Due to the limitations related to my account, the number of places per neighborhood parameter should be set to 100 and the radius parameter to 500.

Summary Libraries:

Pandas: For creating and manipulating dataframes.
Numpy: is a math library to work with N-dimensional arrays.
Folium: Python visualization library would be used to visualize the neighborhoods cluster distribution of using interactive leaflet map.
Scikit Learn: is a collection of numerical algorithms and domain specific toolboxes in this case for importing k-means clustering.
JSON: Library to handle JSON files.
Geopy: To retrieve Location Data
Requests: Library to handle http requests
Matplotlib: Very popular plotting package that provides 2D plotting, as well as 3D plotting module