Final Report | Capstone Project – The Battle of Neighborhoods Finding a Better Place in North York, Toronto

Gaurav Samant

1. Introduction

The purpose of this Project is to help people in exploring better facilities around their neighborhood. It will help people making smart and efficient decision on selecting great neighborhood out of numbers of other neighborhoods in North York, Toranto.

This Project aim to create an analysis of features for a people migrating to North York to search a best neighborhood as a comparative analysis between neighborhoods. The features include median housing price and better school according to ratings, crime rates of that particular area, road connectivity, weather conditions, good management for emergency, water resources both freash and waste water and excrement conveyed in sewers and recreational facilities.

2.Data Section

Data Link: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

Foursquare AP

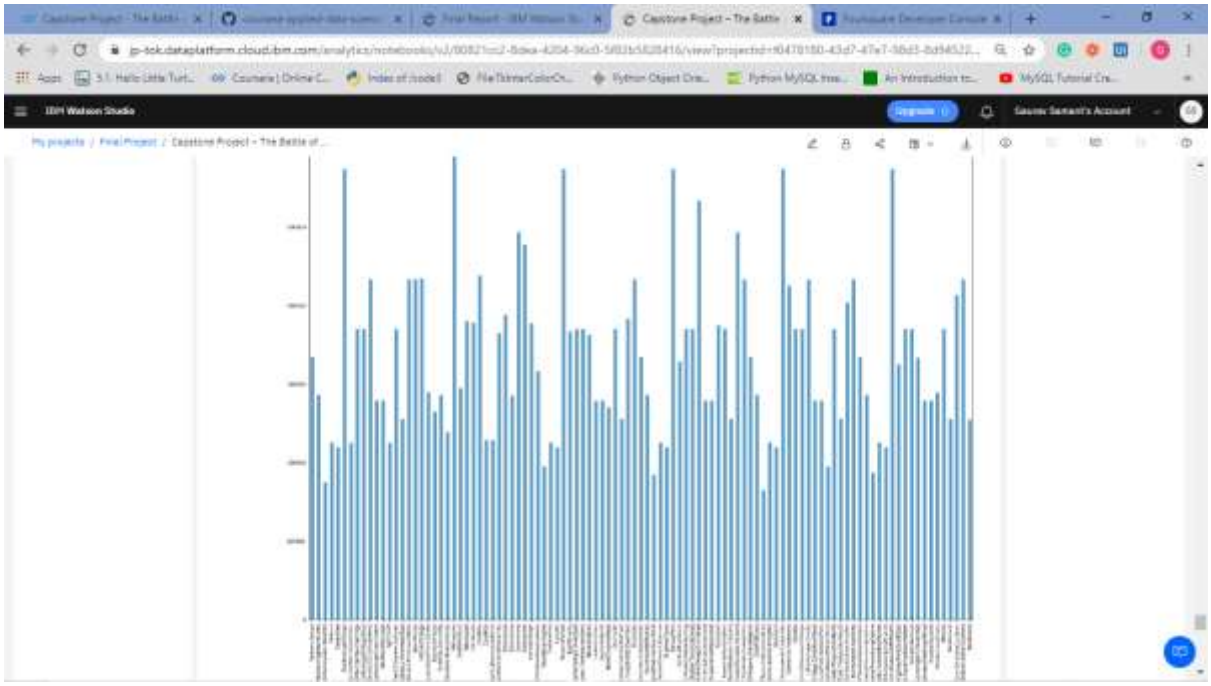
3. Methodology Section

Clustering Approach: To compare the similarities of two cities, we decided to explore neighborhoods, segment them, and group them into clusters to find similar neighborhoods in a big city like New York and Toronto. To be able to do that, we need to cluster data which is a form of unsupervised machine learning: k-means clustering algorithm.

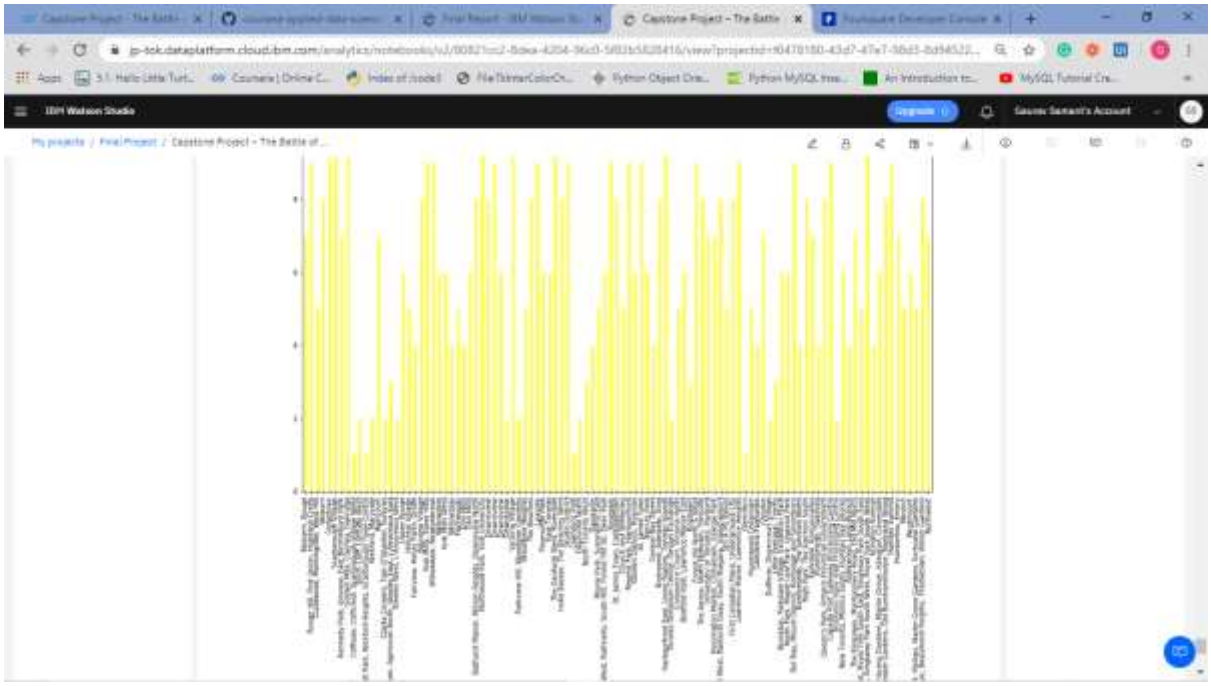
Work Flow: Using credentials of Foursquare API features of near-by places of the neighborhoods would be mined. Due to http request limitations the number of places per neighborhood parameter would reasonably be set .

4. Results Section

Average Housing Price by Clusters in North York



School Ratings by Clusters in North York



Foursquare API:

This project have used Four-square API as its prime data gathering source as it has a database of millions of places, especially their places API which provides the ability to perform location search, location sharing and details about a business.

5. Discussion Section

Problem Which Tried to Solve:

The major purpose of this project, is to suggest a better neighborhood in a new city for the person who are shiffting there. Social presence in society. Connectivity to the airport, bus stand, city center, markets and other daily needs things nearby.

1. Sorted list of house in terms of housing prices
2. Sorted list of schools in

6. Conclusion Section

In this project, using k-means cluster algorithm. Using the charts above results presented to a particular neighborhood based on average house prices and school rating have been made.

The mapping with Folium is a very powerful technique to consolidate information and make the analysis and decision better with confidence.

Libraries Which are Used to Develope the Project:

Pandas

Folium.

JSON

XML

Geocoder

Beautiful Soup

Matplotlib