

Battle of Neighbourhoods



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 Scarborough, Toronto.

Introduction

Lots of people are migrating to various states of Canada and needed lots of research for good housing prices and reputed schools for their children. This project is for those people who are looking for better neighborhoods. For ease of accessing to Cafe, School, Super market, medical shops, grocery shops, mall, theater, hospital, like minded people, etc.

Data

Data Link:

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

Will use Scarborough data-set which we scrapped from Wikipedia on Week 3. Data-set consisting of latitude and longitude, zip codes.

Foursquare API

We will need data about different venues in different neighborhoods of that specific borough. In order to gain that information we will use "Foursquare" location information. Foursquare is a location data provider with information about all manner of venues and events within an area of interest. Such information includes venue names, locations, menus and even photos. As such, the foursquare location platform will be used as the sole data source since all the stated required information can be obtained through the API.

Data Description

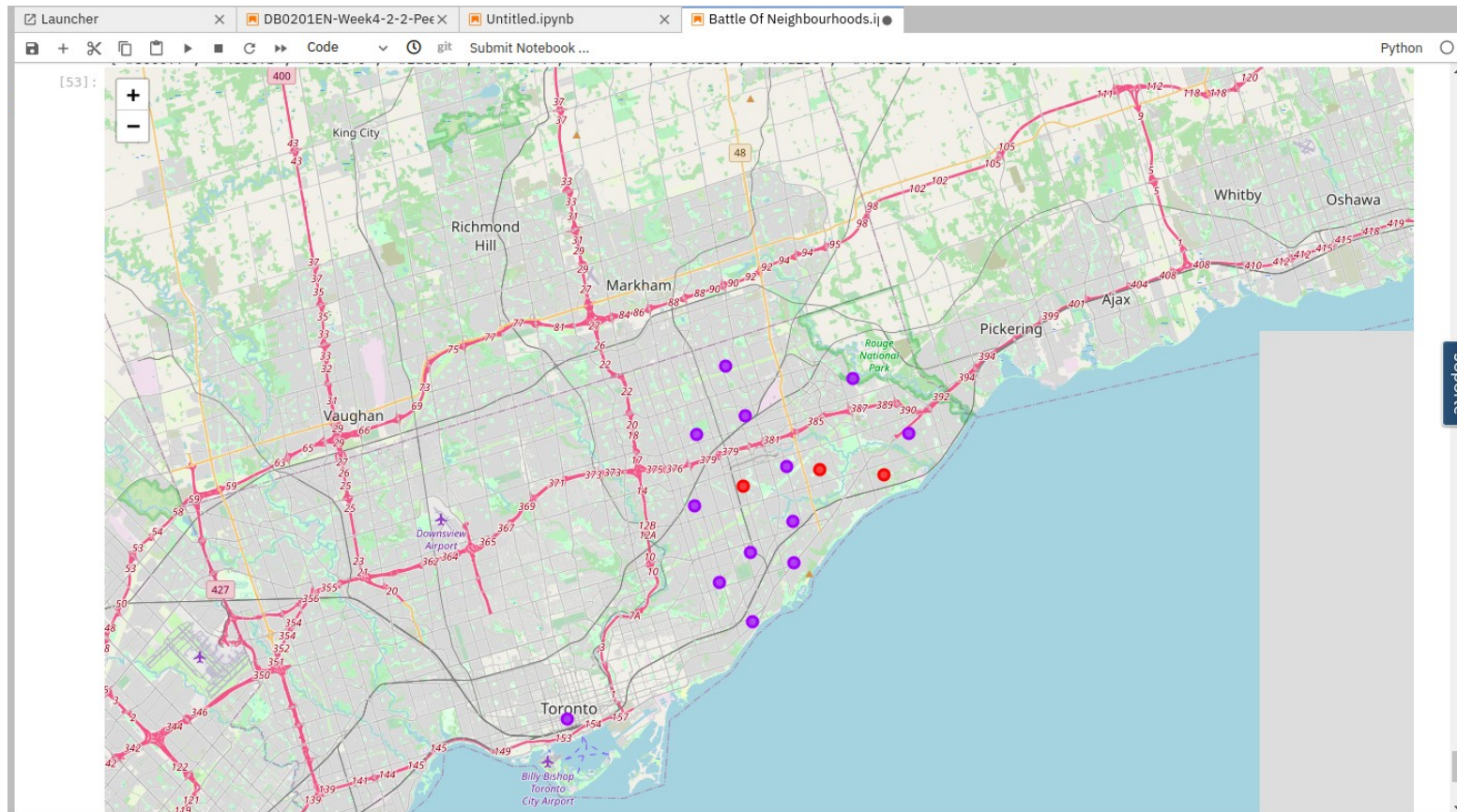
The data retrieved from Foursquare contained information of venues within a specified distance of the longitude and latitude of the postcodes. The information obtained per venue as follows:

1. Neighborhood
2. Neighborhood Latitude
3. Neighborhood Longitude
4. Venue
5. Name of the venue e.g. the name of a store or restaurant
6. Venue Latitude
7. Venue Longitude
8. Venue Category

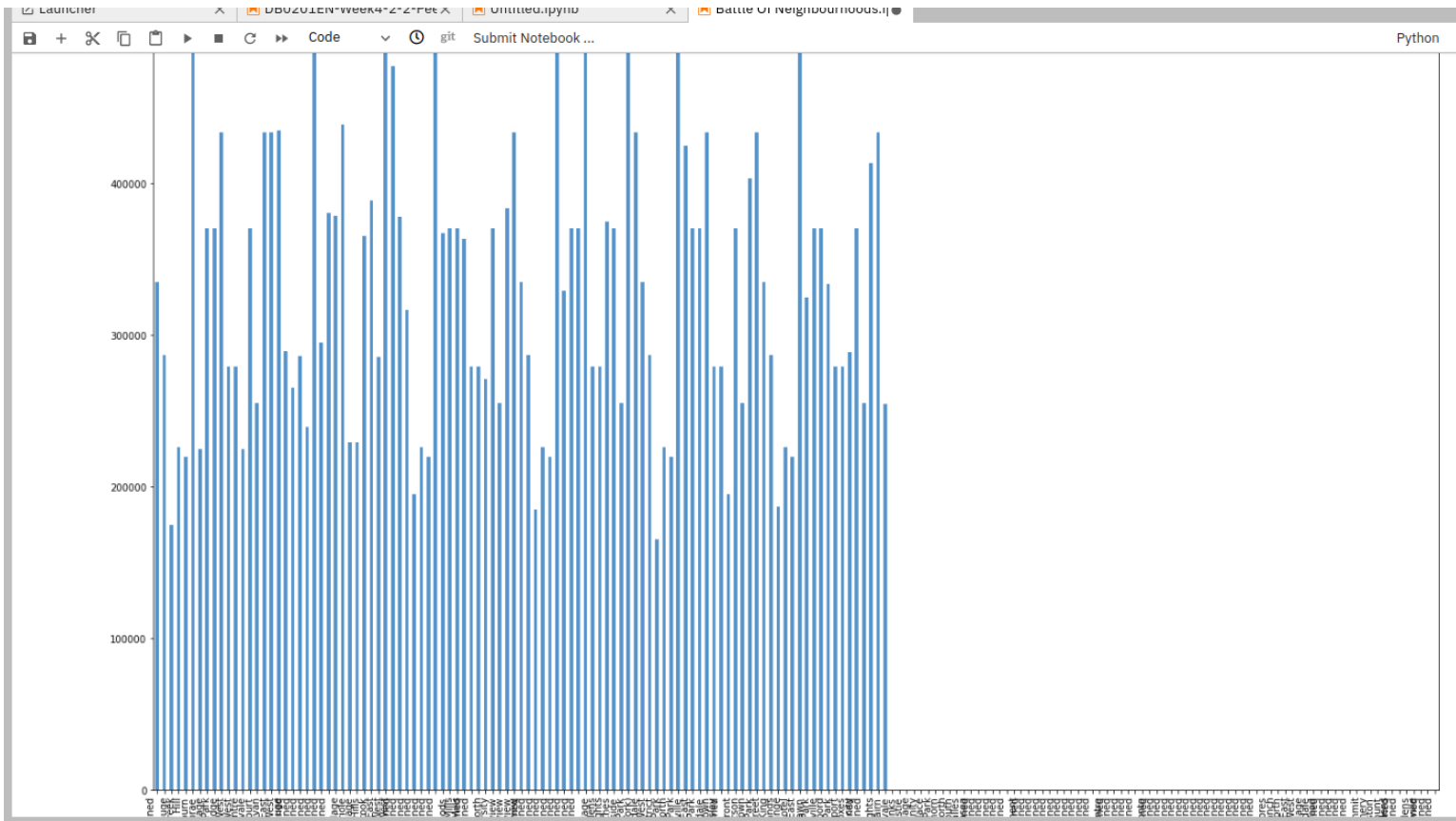
USING K-MEANS

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.

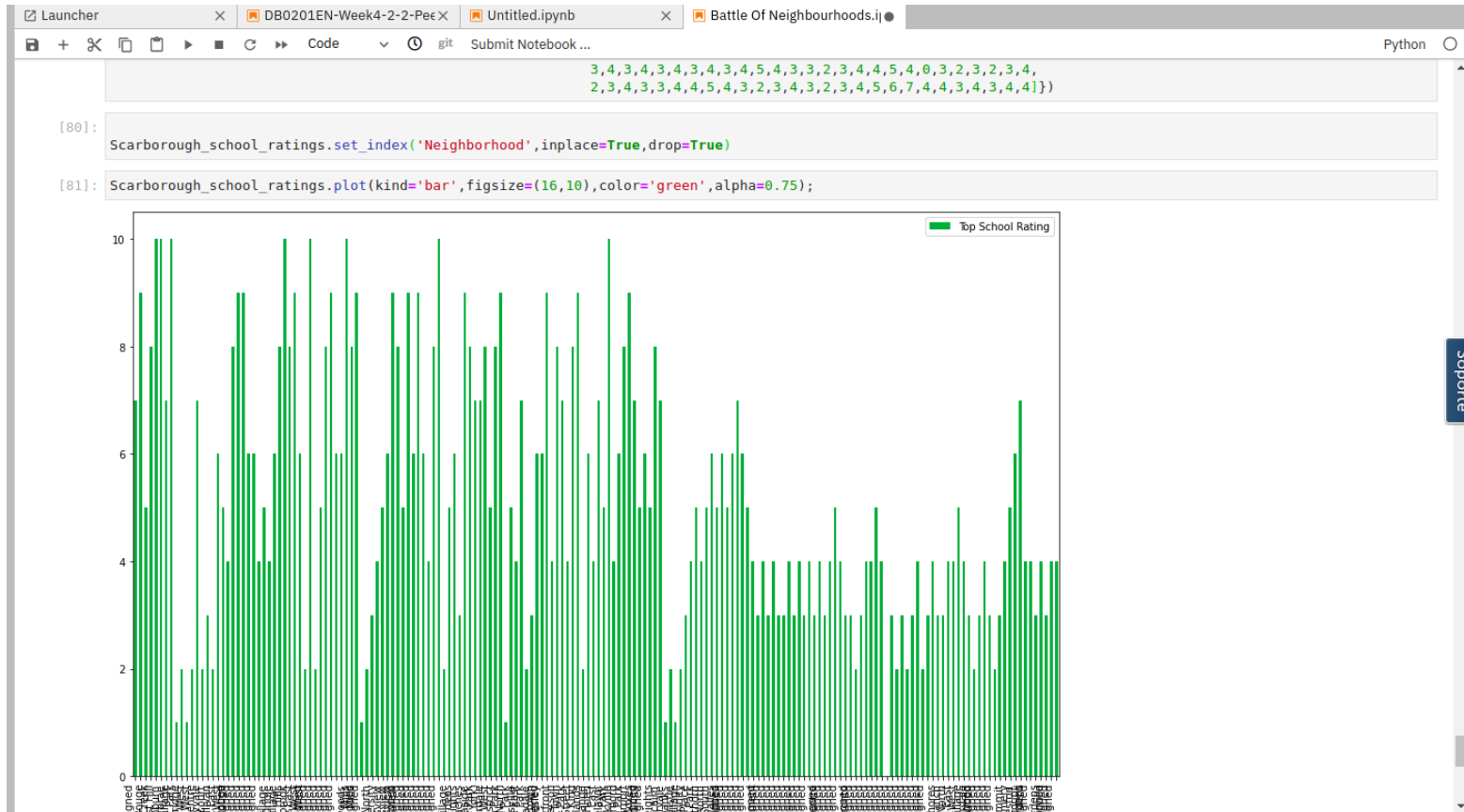
Map with 10 Clusters



Avg house pricing



Ranking of schools



Conclusions

In this project, using k-means cluster algorithm I separated the neighborhood into 10(Ten) different clusters and for 103 different latitude and longitude from data-set, which have very-similar neighborhoods around them. Using the charts above results presented to a particular neighborhood based on average house prices and school rating have been made.

Conclusions

I feel rewarded with the efforts and believe this course with all the topics covered is well worthy of appreciation. This project has shown me a practical application to resolve a real situation that has impacting personal and financial impact using Data Science tools. The mapping with Folium is a very powerful technique to consolidate information and make the analysis and decision better with confidence.