THE BATTLE OF NEIGHBOURHOODS

By Shailesh Thakur

Dec' 2020

INTRODUCTION

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

◦ Sorted list of house in terms of housing prices in ascending or descending order

◦ Sorted list of schools in terms of location, fees, rating and reviews

DATA ACQUISITION & CLEANING

◦ Canada Postal Codes data scrapped from Wikipedia through beautiful soup List_of_postal_codes_of_Canada

◦ Location & venue data from Foursquare API v2 Foursquare

◦ Foursquare is a location data provider with information about all manner of venues and events within an area of interest

◦ For each neighborhood, we have chosen the radius to be 100 meter.

◦ 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 are Neighborhood, Neighborhood Latitude, Neighborhood Longitude, Venue, Name of the venue, Venue Latitude, Venue Longitude
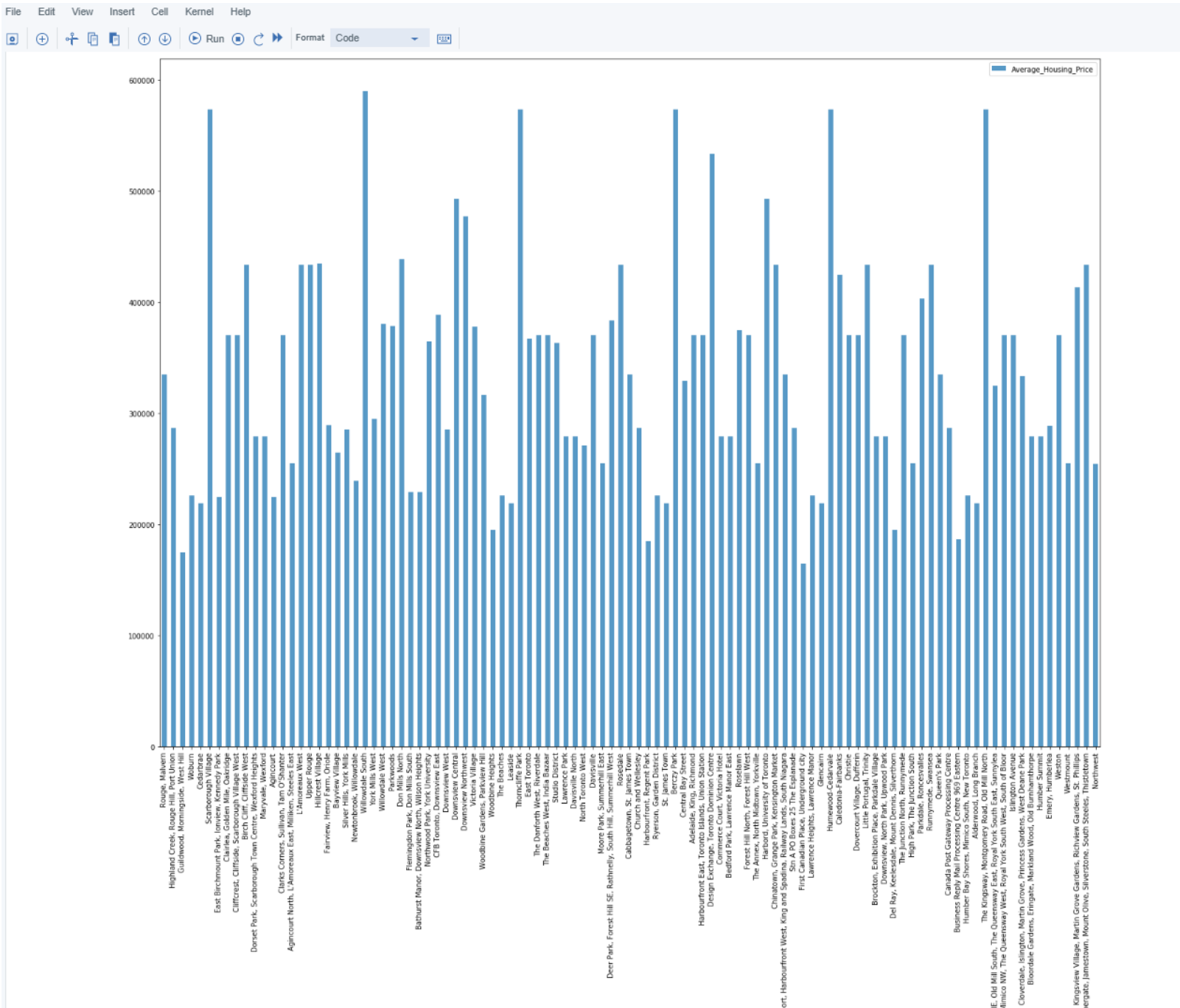
METHODOLOGY

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.
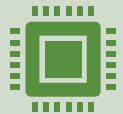
To be able to do that, we need to cluster data which is a form of unsupervised machine learning: k-means clustering algorithm

RESULTS

Map of Scarborough

Separated the neighborhood into 10 different clusters and for 103 different latitude and longitude from dataset, which have very-similar neighborhoods around them.

Average Housing
Price by Clusters
in Scarborough

charts beside shows
neighborhood based on
average house prices.

School Ratings by Clusters in Scarborough

charts beside shows neighborhood based on school rating have been made.

Conclusion

◦ In this Capstone project, using k-means cluster algorithm I separated the neighborhood into 10(Ten) different clusters and for 103 different latitude and longitude from dataset, which have very-similar neighborhoods around them.

◦ Using the charts 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.

◦ Scarborough is a popular destination for new immigrants in Canada to reside.