Predicting The Ideal houses at the desired locality on customer's request

Yashasvi Pamu

June 27th, 2019

1. Introduction

1.1 Background

Finding an ideal home with all the amenities you need has always been an important necessity since time immemorial. People have been compromising their requirements near their homes for other requirements. And as the generation has advanced we can find ideal homes through online and even check them online! The problem is a description of finding an ideal home for my cousin's family who are planning on moving from downtown Toronto to Scarborough.

1.2 Problem

Finding an ideal place to move in based on the customer's needs and requirements or specifications, which may include venues like

restaurants or parks or stores nearby and in the particular neighbourhood.

2 .Data acquisition and cleaning

2.1 Data sources

Data can be obtained through Wikipedia as it contains all the postal codes neighbourhoods as well as details of cities and everything. The data source for this particular problem here is https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

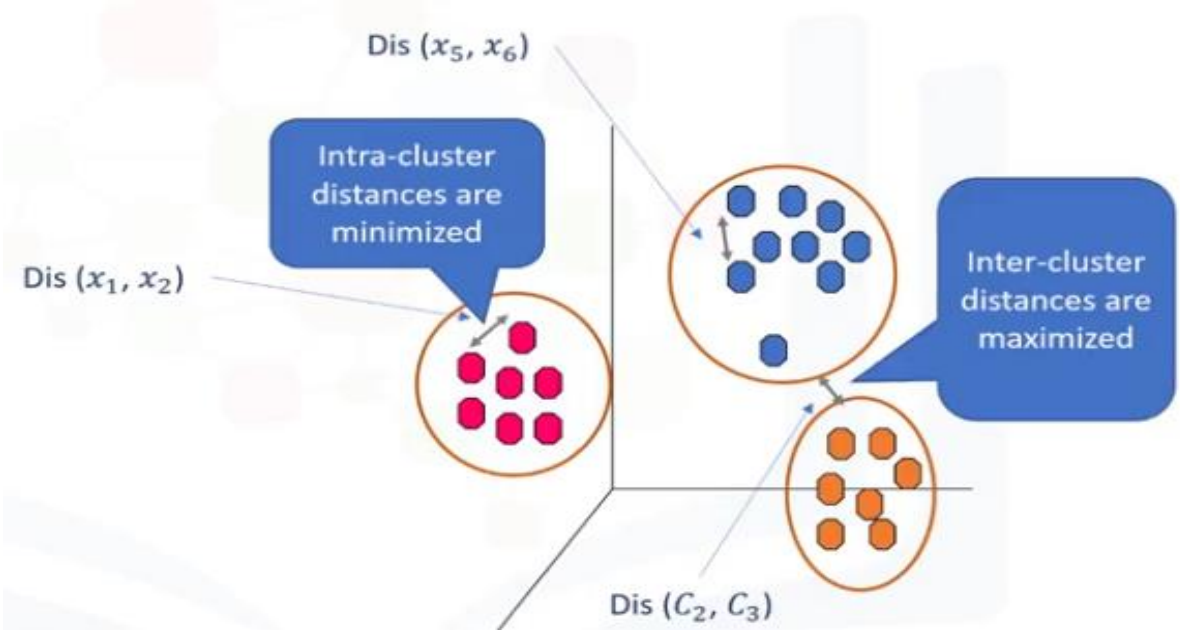
2.2 Data cleaning

Data downloaded or scraped from Wikipedia is converted into one table. There were missing values in Borough of the table which were dropped and the remaining values were taken into consideration. They were then sorted out and the selected neighbourhood values were separated and then were again sorted and distinguished and removed and finally the place is found out by using ML algorithms.

Map of neighbourhoods in Toronto and Scarborough

3 .Predictive Modelling

The model used here is K-means algorithm. K-Means is a type of partitioning clustering, that is, it divides the data into K non-overlapping subsets or clusters without any cluster internal structure or labels. This means, it's an unsupervised algorithm.



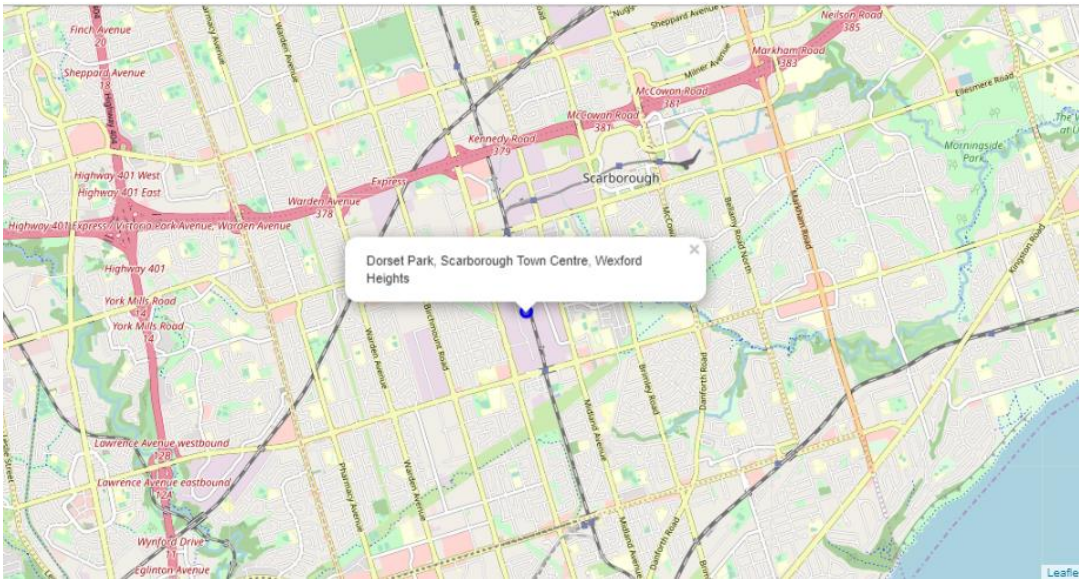
K-means Clustering

4 .Methodology

K-means is used to cluster the data of neighbourhoods according to the requirements mentioned and then geocoder package is installed to provide with latitudes and longitudes and then folium is used to create maps. Then live venue details are gathered from Foursquare API which can be accessed through requests. Then the venues are sorted according to the neighbourhoods and the venues and their ratings and tips provided to them by users or customers who have been there.

5 .Results

After all the processes and sorting and finding the ideal locality in neighbourhoods based on the venues requested by the customer here, an Indian restaurant and an Asian grocery store. And the ideal locality for the family to shift is



Dorset Park, Scarborough Town Centre, Wexford Heights

6 .Conclusion

This is the application of the data science course which I have learnt in the previous 8 courses which helped me predict a place according to the customer requirement which can help me in the real-estate business.