

The battle of Neighborhoods to open a Piano Bar

By P Vamsi Sai Maitreya

Background

Since lots of people nowadays find interest in music and parties to get a take away for sometimes from their regular life. So, if a borough with most of the neighborhoods can be selected for the piano bar if it does not exist there, then it can be really beneficial from business perspective as well as for public convenience.

Problem and Interest

In this project we will try to find an optimal location for a Bar. Specifically, this report will be targeted to stakeholders interested in opening a Piano Bar in a Borough of Toronto, Ontario, Canada which has most neighborhoods.

Since there are lots of bars in Totonto we will try to detect locations or clusters that are not already crowded with bars. We are also particularly interested in areas with no Piano Bars in vicinity.

This report will be targeted to stakeholders interested in opening a Piano Bar in a Borough of Toronto, Ontario, Canada which has most neighborhoods.

Data acquisition

Based on definition of our problem, factors that will influence our decision are:

number of existing bars in the neighborhood (any type of bar)

number of Piano bars in the neighborhood, if any

Following data sources will be needed to extract/generate the required information:

Data of Postal code with Borough and Neighborhoods can be scraped or used as

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

For coordinates of each neighborhood, we can get the data

from https://cocl.us/Geospatial_data

number of bars and their type and location in every neighborhood will be obtained

using **Foursquare API**

coordinate of Toronto center will be obtained using geopy

Data cleaning

Data with Toronto Boroughs can be taken as CSV file and then imported in pandas dataframe. Then all the rows having 'Not Assigned' values should be analyzed and treated accordingly.

Name/Rename all the columns correctly. Data for the latitude and longitude of all the postal codes can be taken from the link mentioned in 'data sources' section and then it should be cleaned properly. Then merge the data of the two dataframes to get the final dataframe to work upon.

Methodology

Now we have the cleaned data to build the model

In this project, first we'll see and explore all the locations in Toronto.

Then we'll find the Borough with most neighborhoods.

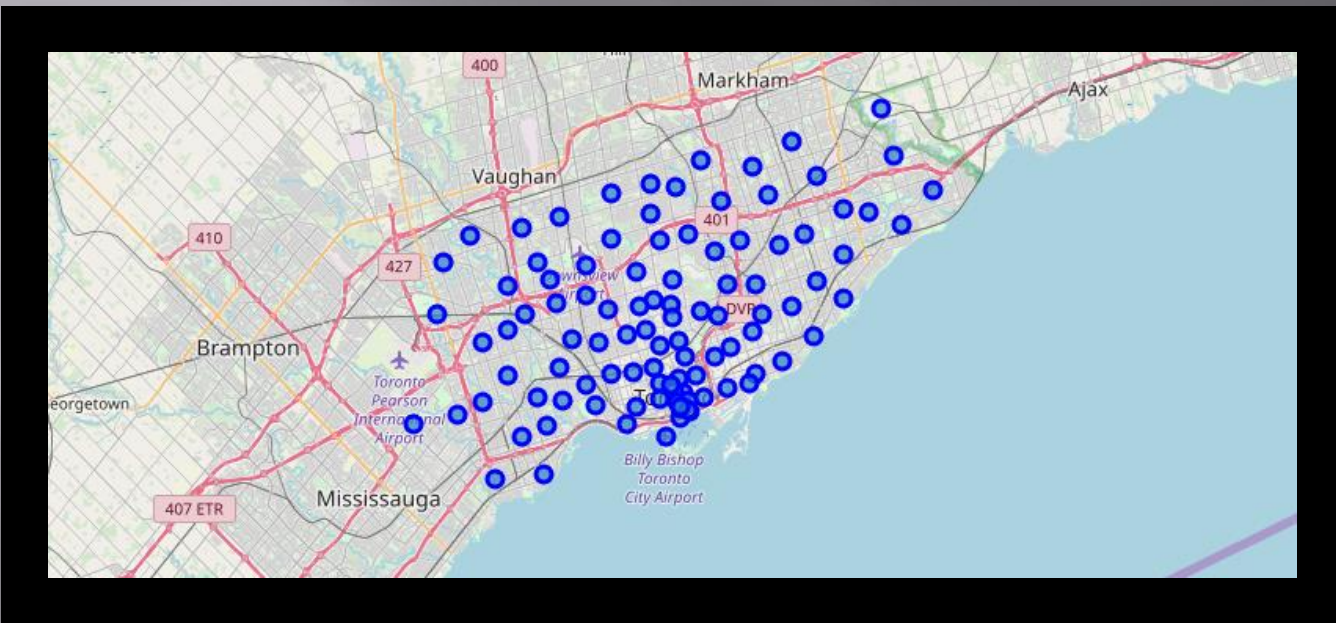
We'll explore all the locations in the Borough.

Then we'll cluster all the locations in that borough.

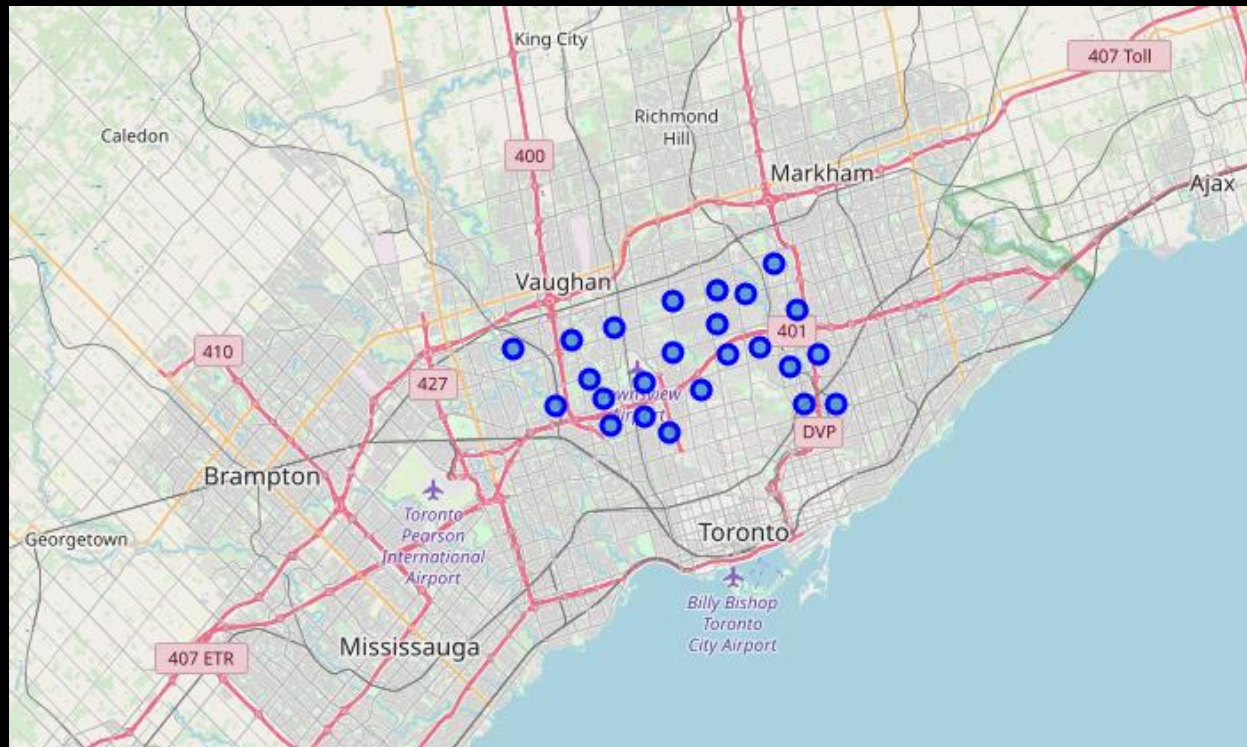
After clustering the neighborhoods in the borough, we'll analyze which cluster has least or no Bar.

Also, we'll focus on the cluster(s) which has no piano bar.

Let's visualize Toronto city and the neighborhoods in it.



Young players improve, old players decline



Foursquare

Next, we are going to start utilizing the Foursquare API to explore the neighborhoods and segment them.

We'll explore the first neighborhood in our dataframe with its name, latitude and longitude and, how many venues were returned by Foursquare. We saw that the number was five. Then we explored Neighborhoods in North York.

We'll first look for number of unique categories which can be curated from all the returned venues and it returned 107.

Later we'll analyze each neighborhood. We see each neighborhood along with the top 5 most common venues. Then we'll apply k -means to cluster the neighborhood into 5 clusters. We will remove all the rows with NAN value.

Finally, we'll visualize the resulting clusters



Results

our **first choice** for opening a Piano bar can be **cluster 4** as it doesn't have any bar and crowded with favorable venues.

Our **second choice** can be **cluster 1** since it very crowded with favorable venues and it doesn't contain any Piano Bar though it contains bars and pubs but, piano bar would definitely attract visitors.

Our **third choice** can be **cluster 5** as it contains airport and few restaurants which can attract people for a piano bar though it's not crowded with many venues.

Conclusion

From the above results and discussion we can conclude that one can go ahead to open a piano bar in cluster 1 as the best cluster. But if from cost perspective any of the cluster 4 or five are better then without a doubt they can go for that. There are good chances of profits opening a piano bar following this strategy.