Business Site Location Analysis by Segmenting Neighborhoods in Ottawa, ON, CANADA

Sher Baz Khan

In this project we analyze data locations by segmenting neighborhoods in the City of Ottawa, Ontario, CANADA. The objective of this analysis is to provide a framework to facilitate the stakeholders who might be interested in locating a suitable area in certain neighborhood of the city to a business venture startup. The analysis presented in this project will only provide a starting point towards finding an optimum location by looking at those areas where a particular business in not established.

1 Introduction: Business Problem

In this project, We analyze the location data by segmenting and clustering neighborhoods in Ottawa, ON - the capital city of CANADA. The objective of this analysis is to venture out feasibility of establishing a small business enterprise in a particular neighborhood in the city of Ottawa. For this purpose, we are focusing on a particular neighborhood called Orleans-Ottawa which is one of the fast growing neighborhoods in the capital city. We are interested in setting up a small business enterprise, say, a Bookstore in Orleans-Ottawa neighborhood. Therefore our objective dig out whether it is feasible to set up a Bookstore in the Orleans neighborhood or not. Our target is to check whether there is such an enterprise already existing in that neighborhood apart from other basic businesses. If there is no such business established yet, then we will of course go for it. In this project we try to find an optimal location for a Bookstore. This report will be targeted to those stakeholders interested in opening a Bookstore in Orleans-Ottawa, ON. Since there are so many bookstores already established in various neighborhoods in Ottawa we will try to detect those neighborhoods where there are no such bookstores established. By using methods of data science we will generate some potential neighborhoods that would be based on this single criterion. We will show different commonality conditions of various neighborhoods which will provide the most optimum location to choose from by interested stakeholders.

2 Importance of Location in a Business Venturing

When it comes to a site selection for a future business start up, location plays fundamentally vital role as it provides the direct link between demand and supply in a given marketplace - demand by customers for services or goods and supply of essential resources. The importance of location varies from one sort of business to another. In the case of a retail outlet, location must depend on the traffic flow, population density, interests of customers and many other factors. While establishing a new outlet in a certain location it must be reviewed and take into account several other locations. It should not be a once-only decision. Business location in necessary for

successful operations and growth. The suitability of a business location and the customers potential in terms of their purchase power may lead a business to succeed or fail.

3 Study Area

Ottawa [1] is the national capital city of Canada. It is situated in the east of southern Ontario, near the city of Montral. The city of Ottawa borders Gatineau of the French speaking Quebec province. Ottawa has a city population of 964,743 and a metropolitan population of 1,323,783 making it the fourth-largest city in Canada. Ottawa has been focus of attractions for traders from Europe and America in the history. It was shaped by construction of the Rideau Canal, a UNESCO World Heritage Site and the lumber industry. Ottawa is known for being welcoming, inclusiveness and diversity of people of different backgrounds.

Ottawa's real development started in the second half of the 20th century. In 1960s, the Greber Plan transformed the capital's appearance and removed much of the old industrial infrastructure. Ottawa became known as Silicon Valley North in 1980s when large high tech companies were established. It brought economic growth and as a consequence caused a rapid increase in population. The city amalgamated all neighboring areas in 2001. Ottawa is continuously progressing in areas such as population, transportation and economic growth [2].

4 Methodology

We focus our attention on those areas where there is no Bookstore established in the neighborhoods of our interest at the time of our investigation. This will be the only criterion to quest for a prospective future business venture. Next, we cluster few neighborhoods including the neighborhood Orleans-Ottawa so that we would be able to analyze neighboring locations keeping the above mentioned criterion in view. We create a data frame using the Pandas library in Python. We import KMeans for clustering the neighborhoods from sklearn.cluster library. Furthermore, from geopy.geocoders we use Nominatim library in order to convert addresses into latitude and longitude values.

We use the Foursquare API [4] to explore various neighborhoods in Ottawa ON. We use the **explore** function to get the most common venue categories in each neighborhood, and then use this feature to group the neighborhoods into clusters. We use the k-means clustering algorithm to complete this task. Finally, we use the Folium library to visualize all neighborhoods in Ottawa and their emerging clusters. We in particular focus on the the neighborhood Orleans-Ottawa to see the possibility of establishing a Bookstore in that neighborhood. Next we show maps of all such locations by creating k-means clustering method in order to specify promising areas so that it could present an optimal venue location to be considered by stakeholders.

```
[2]: data = pd.read_html('https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_K')

     df = pd.DataFrame(data[0])
     df.head()
```

[2]:

	0	1	2	3	4	5
0	K1AGovernment of CanadaOttawa and Gatineau off...	K2AOttawa(Highland Park / McKellar Park /Westb...	K4AOttawa(Fallingbrook)	K6AHawkesbury	K7ASmiths Falls	K8APembrokeCentral and northern subdivisions
1	K1BOttawa(Blackburn Hamlet / Pine View / Sheff...	K2BOttawa(Britannia /Whitehaven / Bayshore / P...	K4BOttawa(Navan)	K6BNot assigned	K7BNot assigned	K8BPembroke(Pleasant View / Fairview)
2	K1COttawa(Orleans)	K2COttawa(Queensway / Copeland Park / Central ...	K4COttawa(Cumberland)	K6CNot assigned	K7CCarleton Place	K8CNot assigned
3	K1EOttawa(Queenswood)	K2EOttawa(Eastern Nepean: Fisher Heights/ Park...	K4ENot assigned	K6ENot assigned	K7ENot assigned	K8ENot assigned
4	K1GOttawa(Riverview / Hawthorne / Canterbury /...	K2GOttawa(Centrepointe, Meadowlands, City View...	K4GNot assigned	K6GNot assigned	K7GGananoque	K8GNot assigned

Figure 1: Web scraping

5 Data Dependencies

In order to achieve our objective we will fetch data using the Google's search engine from the wikipedia's webpage which contains all the neighborhoods in the City of Ottawa-ON. This data is in raw form. We observed that the data does not include GPS coordinates corresponding to the postal codes of the neighborhoods. To overcome this issue we extract the data for respective coordinates from another web source [4].First and foremost, we will get rid of that part of data which is not related to our investigation like removing those neighborhoods of which postal codes have not yet been assigned or not in use. Before manipulating the data it is necessary to cleanse it. We then incorporate the Latitudes and Longitudes corresponding to the postal codes of all neighborhoods.

6 Web Scraping

First of all we require a dataset in order to segment the neighborhoods which contains the location data like, postal codes of respective neighborhoods as well as Latitudes and Longitudes of corresponding neighborhoods. For this, we download the dataset from the web source and explore the data. As we see above that this dataset is in the raw form. It does not contain the *Latitude* and *Longitude* values. This needs to be refined and be cleansed. In order to accomplish this we save this data into a csv file and incorporate the Latitude and Longitude values corresponding to the postal codes. We use another web source https://www.gps-coordinates.net to get these values. We then upload the updated csv file named final project.csv. Then we transform this data into a pandas dataframe.

7 Foursquare API

We now start segmenting the neighborhoods and explore using the Foursquare API. We get the request get-category-type to get the information in the items key from the

[3]:
```
data = pd.read_csv("final_project.csv")
df = pd.DataFrame(data)
df.head()
```

[3]:

	PostalCode	Neighborhoods	Latitude	Longitude
0	K1B	Blackburn Hamlet / Pine View / Sheffield Glen	45.42042	-75.59603
1	K1C	Orleans	45.46253	-75.53009
2	K1E	Queenswood	45.47389	-75.50540
3	K1G	Riverview / Hawthorne / Canterbury / Hunt Club...	45.38954	-75.62517
4	K1H	Alta Vista / Billings Bridge	45.38777	-75.65906

Figure 2: Uploading the csv file

```
# clean columns
nearby_venues.columns = [col.split(".")[-1] for col in nearby_venues.columns]

nearby_venues.head()
```

[14]:

	name	categories	lat	lng
0	Rexall Drugstore	Pharmacy	45.461416	-75.525040
1	Boomerang Kids	Thrift / Vintage Store	45.461037	-75.524897
2	Louis Perrault Park	Playground	45.458599	-75.530604
3	Cedar Valley Lebanese Food	Restaurant	45.461094	-75.524313

Figure 3: Exploring the location data

Foursquare API. We also cleaned the *json* and structure it into a pandas dataframe.

8 Explore all Neighborhoods in and around Orleans-Ottawa

Now write the code to run the above function on each neighborhood and create a new dataframe called venues dataframe. We see the size of the resulting dataframe. We found out how many venues were returned for each neighborhood.

9 Analysis of Neighborhoods

We use the following code to start analyzing neighborhoods. Let's look at top 5 most common venues of each neighborhood.

Let's display the top 10 venues for each neighborhood.

10 Clustering Neighborhoods

Here we now run k-means to cluster the neighborhoods into 5 clusters. We now create a new dataframe showing top 10 venues for the Orleans-Ottawa neighborhood that

```
[20]: print(df_venues.shape)
      df_venues.head()
```

(527, 7)

[20]:

	Neighborhoods	Neighborhoods Latitude	Neighborhoods Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Blackburn Hamlet / Pine View / Sheffield Glen	45.42042	-75.59603	Apple Saddlery	45.419272	-75.599590	Shoe Store
1	Blackburn Hamlet / Pine View / Sheffield Glen	45.42042	-75.59603	Eclipse Asian Cuisine	45.418327	-75.596766	Asian Restaurant
2	Blackburn Hamlet / Pine View / Sheffield Glen	45.42042	-75.59603	Sushi Kan	45.418031	-75.597849	Sushi Restaurant
3	Blackburn Hamlet / Pine View / Sheffield Glen	45.42042	-75.59603	Big Al's Aquarium Services	45.418138	-75.595990	Pet Store
4	Blackburn Hamlet / Pine View / Sheffield Glen	45.42042	-75.59603	Marks	45.418654	-75.595985	Clothing Store

Figure 4: Exploring venues within neighborhoods

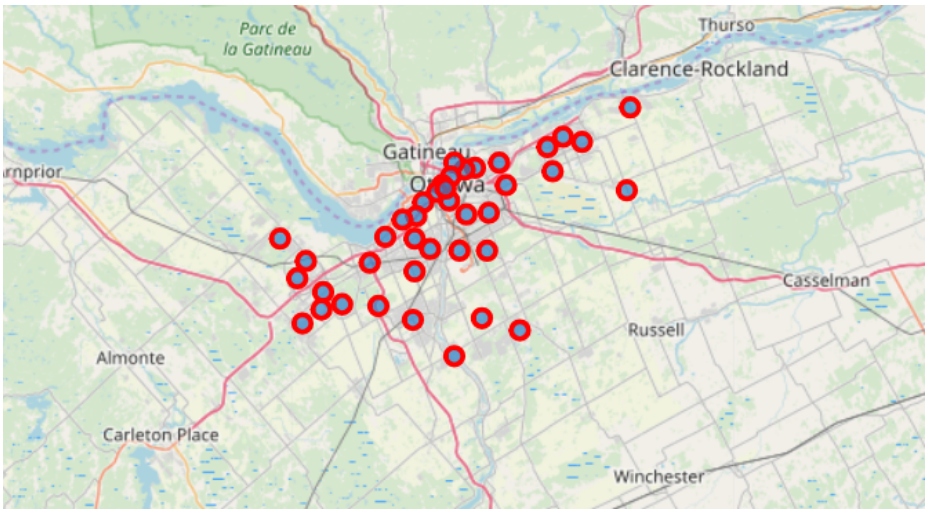

Figure 5: A map of Ottawa-ON with neighborhoods superimposed on top

```
[21]: df_venues.groupby('Neighborhoods').count()
```

[21]:

Neighborhoods	Neighborhoods Latitude	Neighborhoods Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Alta Vista / Billings Bridge	3	3	3	3	3	3
Barrhaven	38	38	38	38	38	38
Beacon Hill / Cyrville / Carson Grove	5	5	5	5	5	5
Beaverbrook / South March	3	3	3	3	3	3
Bells Corners / Arlington Woods/Redwood / Qualicum / Crystal Beach	4	4	4	4	4	4
Blackburn Hamlet / Pine View / Sheffield Glen	11	11	11	11	11	11
Blossom Park / Greenboro / Leitrim) / Findlay Creek	11	11	11	11	11	11
Bridlewood	4	4	4	4	4	4

Figure 6: Collecting the venues from each neighborhood

[23]:

```
df_onehot = pd.get_dummies(df_venues[['Venue Category']], prefix="", prefix_sep="")

# add neighborhood column back to dataframe
df_onehot['Neighborhoods'] = df_venues['Neighborhoods']

# move neighborhood column to the first column
fixed_columns = [df_onehot.columns[-1]] + list(df_onehot.columns[:-1])
df_onehot = df_onehot[fixed_columns]

df_onehot.head()
```

[23]:

	Neighborhoods	Adult Boutique	Airport Terminal	American Restaurant	Art Gallery	Asian Restaurant	Athletics & Sports	Auto Dealership	Auto Workshop	BBQ Joint	Bagel Shop	Bake
0	Blackburn Hamlet / Pine View / Sheffield Glen	0	0	0	0	0	0	0	0	0	0	
1	Blackburn Hamlet / Pine View / Sheffield Glen	0	0	0	0	1	0	0	0	0	0	
	Blackburn											

Figure 7: Collecting unique categories from all the returned venues

```
----Alta Vista / Billings Bridge----
                  venue  freq
0         Moving Target  0.33
1            Playground  0.33
2  Fast Food Restaurant  0.33
3     Other Repair Shop  0.00
4         National Park  0.00


----Barrhaven----
                  venue  freq
0  Fast Food Restaurant  0.11
1            Restaurant  0.08
2         Grocery Store  0.08
3           Coffee Shop  0.08
4   American Restaurant  0.05


----Beacon Hill / Cyrville / Carson Grove----
                   venue  freq
0  Entertainment Service   0.2
1                  Café   0.2
2         Sandwich Place   0.2
3     Mexican Restaurant   0.2
4       Malay Restaurant   0.2
```

Figure 8: Top 5 most common venues

```
----Blackburn Hamlet / Pine View / Sheffield Glen----
              venue  freq
0       Golf Course  0.09
1  Asian Restaurant  0.09
2         Pet Store  0.09
3  Sushi Restaurant  0.09
4              Park  0.09


----Blossom Park / Greenboro / Leitrim) / Findlay Creek----
                  venue  freq
0      Department Store  0.18
1  Fast Food Restaurant  0.09
2        Discount Store  0.09
3     Convenience Store  0.09
4         Auto Workshop  0.09


----Bridlewood----
                      venue  freq
0                      Park  0.50
1        Convenience Store  0.25
2              Bus Station  0.25
3     Outdoor Supply Store  0.00
4  New American Restaurant  0.00
```

Figure 9: Top 5 most common venues cont...

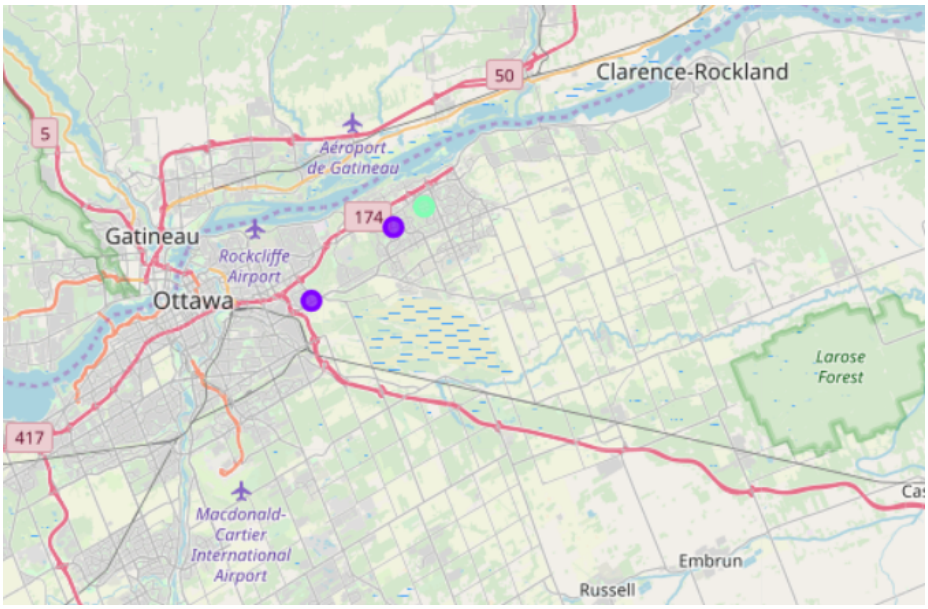

Figure 10: A cluster of Orleans neighborhoods

[29]:

	Neighborhoods	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
0	Alta Vista / Billings Bridge	Playground	Fast Food Restaurant	Moving Target	Convenience Store	Farmers Market	Event Space	Ethiopian Restaurant	Entertainment Service	English Restaurant
1	Barrhaven	Fast Food Restaurant	Coffee Shop	Restaurant	Grocery Store	Pizza Place	American Restaurant	Pet Store	Department Store	Hardware Store
2	Beacon Hill / Cyrville / Carson Grove	Mexican Restaurant	Malay Restaurant	Café	Sandwich Place	Entertainment Service	Yoga Studio	Electronics Store	Farmers Market	Event Space
3	Beaverbrook / South March	Fast Food Restaurant	Coffee Shop	Café	Yoga Studio	Electronics Store	Farmers Market	Event Space	Ethiopian Restaurant	Entertainment Service
4	Bells Corners / Arlington Woods/Redwood / Qual...	Grocery Store	Pharmacy	Clothing Store	Liquor Store	Dog Run	Farmers Market	Event Space	Ethiopian Restaurant	Entertainment Service

Figure 11: Top 10 venues for each neighborhood

```
df_merged.head() # check the last columns!
```

[37]:

	PostalCode	Neighborhoods	Latitude	Longitude	Cluster Labels	ClusterLabels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	K1B	Blackburn Hamlet / Pine View / Sheffield Glen	45.42042	-75.59603	1	1	Food Truck	Clothing Store	Furniture / Home Store	Sushi Restaurant	Shoe Store
1	K1C	Orleans	45.46253	-75.53009	1	1	Playground	Pharmacy	Restaurant	Thrift / Vintage Store	Dive Bar
2	K1E	Queenswood	45.47389	-75.50540	3	3	Park	Home Service	Dog Run	Fast Food Restaurant	Farmers Market

Figure 12: Creating a new dataframe that includes Orleans neighborhood

includes the cluster. We now create a new dataframe showing top 10 venues for the Orleans-Ottawa neighborhood that includes the cluster.

11 Results and Discussion

We have found that many Bookstores are already set up in various neighborhoods in the city of Ottawa, ON. However, there are locations in certain neighborhoods where there is a shortage of such business ventures. There seem to be higher chances to setup a new business like a Bookstore when there is acute need of one, keeping some other vital factors in mind as mentioned above. It is clearly seen from the above analysis that in the neighborhood of our interest i.e., Orleans-Ottawa where there is no such business (Bookstore) established at the time of our investigation.

It shows that the neighborhood of our interest provides an optimum location for stakeholders in their quest for investment. It is worth mentioning here that though there are other vital factors where a prospective investor should look into apart from the scarcity of similar businesses in a particular location or neighborhood. For a prospective business owner there must be more than one locations available to choose the best one. It is therefore, we have mentioned in our analysis more than one neighborhood apart from Orleans-Ottawa so that a stakeholder should have more than one options available to choose the best location.

It was our interest in this project that we could be able to identify that some locations exit in certain neighborhoods where one can investigate in search for a future

business venture. But this does not mean that those are the only optimal locations for setting up a new Bookstore! The objective of our analysis was to provide a first hand information on exploring areas in the city of Ottawa with either existing or non-existing Bookstores in various neighborhoods. It is to be noted that the locations identified in our analysis should be considered as a starting point only. Therefore, it is recommended that factors like customers' participation in the prospective business, accessibility, growth, and many other relevant factors and conditions must also be taken into account.

12 Conclusion

Our objective in this project was to explore various areas of Ottawa-Ontario in order to facilitate stakeholders somehow in searching out options for optimal locations for establishing a new bookstore. Using the Foursquare data analysis we identified neighborhoods to generate a collection of locations which satisfy our basic and sole requirement regarding the existence of bookstores in various neighborhoods of the city. Potential locations were clustered in order to create neighborhoods of interest so that exploration may be carried out by stakeholders. Options in this direction were made available to select from so that interested stakeholders may find it easy to choose in view of other relevant conditions. It will provide a ready-made choices for interested investors to have an estimated decision based on other vital characteristics of locations in certain neighborhoods.

13 Future work

We know that the City of Ottawa is the corporate entity of municipal government in Ottawa, Ontario, Canada. This corporation is responsible for providing services to the public. Administratively, Ottawa is composed of 23 wards. Each ward is represented by city councilor and a mayor. In order to extend the analysis presented in this project one can obtain a new dataset containing all the neighborhoods in conjunction with all the 23 wards in the city of Ottawa. By doing this it can open further avenues of in-depth analysis so that stakeholders can diversify their search in the quest for more suitable locations by providing more available options.

References

[1] Wikipedia

[2] Ottawa.ca

[3] gps-Coordinates.net

[4] Forsquare API