CAPSTONE PROJECT - The Battle of Neighborhoods

Where Tim Hortons should open its first shop in Los Angeles

Ahnaf Abid Hassan

Introduction

Background

Tim Hortons is Canada's largest and most popular fast-food restaurant with 4300 shops in Canada. [1] The brand is known for its coffee and baked goods. Although the brand is popular in Canada, it has not found the same level of success in the United States. With about 730 shops there and the majority opened in New York, the company has declining average sales every year. [2]

Problem

To increase their sales, Tim Hortons should open shops in Los Angeles. Los Angeles is the second-most populous city in the United States, and coffee shops are very popular there.[3] Tim Hortons needs to locate the shops in busy neighborhoods where coffee shops are frequently visited in order for their sales and profits to increase. This project aims to discover the neighborhood in Los Angeles where Tim Hortons should open their first shop.

Interest

Since this project is about where Tim Hortons should expand in Los Angeles, it will be of interest to the company and its stakeholders. It should also be of interest to Starbucks, McDonald's, Burger King and Dunkin Donuts as they are Tim Hortons' biggest competitors. [4]

Data

4 data sources will be used in this project to discover the neighborhood in Los Angeles where Tim Hortons should open its first store. Detail description of the data and its usage is below.

1. Neighborhoods in Los Angeles

This dataset was scraped from the *Los Angeles Times* website. It contains the names of all the 272 neighborhoods in the Los Angeles County. The names are important in order to find further details about each neighborhood.

	Name	Region
0	Acton	Antelope Valley
1	Adams-Normandie	South L.A.
2	Agoura Hills	Santa Monica Mountains
3	Agua Dulce	Northwest County
4	Alhambra	San Gabriel Valley
5	Alondra Park	South Bay
6	Altadena	Verdugos
7	Angeles Crest	Angeles Forest
8	Arcadia	San Gabriel Valley
9	Arleta	San Fernando Valley
10	Arlington Heights	Central L.A.
11	Artesia	Southeast

Figure 1. Some of the neighborhoods in Los Angeles

2. OpenCage Geo Code

OpenCage GeoCode was used to generate the geospatial coordinates of all the neighborhoods in Los Angeles. These coordinates, neighborhood names and the radius from the coordinates are required by Foursquare location data to know which venues are popular in the neighborhoods.

	City	Region	Latitude	Longitude
0	Acton	Antelope Valley	34.480742	-118.186838
1	Adams-Normandie	South L.A.	34.033081	-118.297115
2	Agoura Hills	Santa Monica Mountains	34.147910	-118.765704
3	Agua Dulce	Northwest County	34.496382	-118.325635
4	Alhambra	San Gabriel Valley	34.093042	-118.127060

Figure 2. Generated coordinates of the neighborhoods

3. Foursquare Location Data

Foursquare location data are required in two instances in this project.

1. To explore each neighborhood and find the most trending venues. Such as parks, coffee shops, shopping mall, etc.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
0	Adams-Normandie	Sushi Restaurant	Café	Taco Place	Donut Shop
1	Agoura Hills	Gym	Pharmacy	Deli / Bodega	Gas Station
2	Agua Dulce	Grocery Store	Pizza Place	Bakery	Café
3	Alhambra	Ice Cream Shop	Bakery	Burger Joint	Café
4	Alondra Park	Gym / Fitness Center	Park	Baseball Field	Fabric Shop

Figure 3. Most common venues in some of the neighborhoods

2. To locate Starbucks' in the neighborhood.

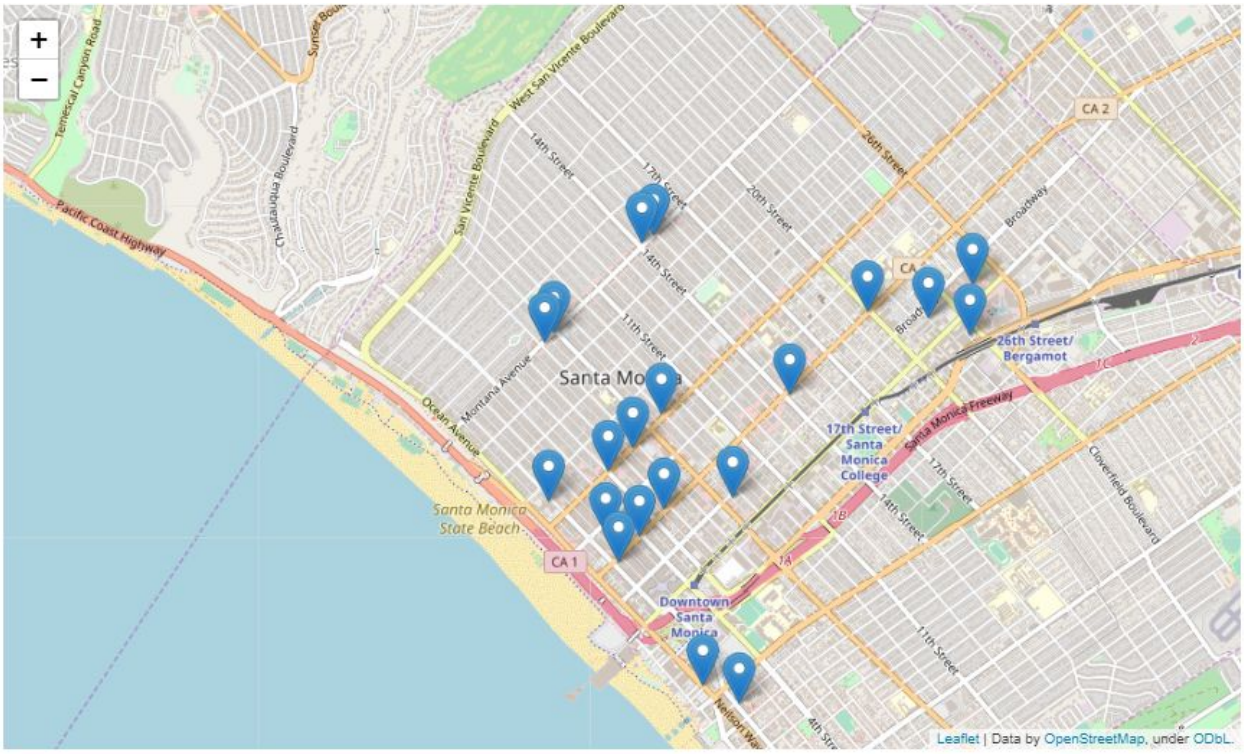

Figure 4. The markers show the location of Starbucks in Santa Monica

4. Population in each Neighborhood

The population data is needed to compare potential customer traffic to coffee shops in the neighborhoods. A higher population of the neighborhood means more potential customers for

the shops. This data was collected from Google and inputted into an excel file. The excel file was read into a data frame using python.

	Neighborhood	Population
0	Atwater Village	29092
1	Century City	8554
2	Claremont	36015
3	Culver City	39283
4	Larchmont	62733
5	Los Feliz	45616
6	Monrovia	37061
7	Norwalk	106084
8	Pacific Palisades	24651
9	Santa Monica	92306
10	Sawtelle	45742
11	South Park	35323
12	Windsor Square	8495

Figure 5. Neighborhoods and their population

Methodology

To find a precise neighborhood in Los Angeles for Tim Hortons to open its first shop the data, mentioned above, were analyzed and modified using Jupyter Notebook. The methods used in the notebook can be broken down and explained in 9 steps.

Step 1 - Finding all Neighborhoods in Los Angeles

The names of all the neighborhoods in Los Angeles was sourced from the *Los Angeles Times* website and converted into a data frame, as shown in Figure 1. The data frame contains 272 rows, 2 columns and 0 null values.

Step 2 - Using OpenCage to generate coordinates of all Neighborhoods

Using the OpenCageGeocode API, each neighborhood latitude and longitude were generated and added to the data frame, shown in Figure 2. The OpenCageGeocode needed the names of the neighborhood and the county name to generate the coordinates.

Step 3 - Visualizing Neighborhoods in Los Angeles

The folium library was used to generate a map of Los Angeles. On that map, the coordinates were used to pinpoint the location of each neighborhood, shown as the blue circles on the map below. The names of the neighborhoods are the labels of the blue circles.

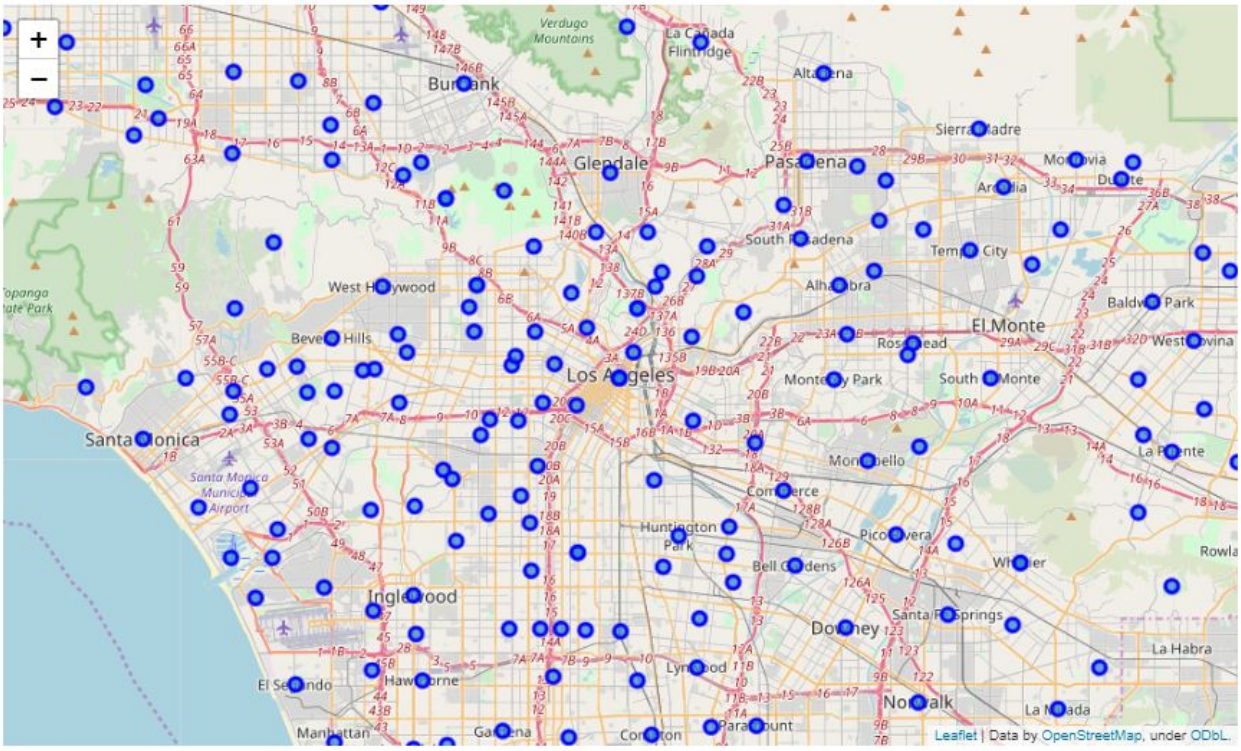

Figure 6. Map of Los Angeles with its neighborhoods

Step 4 - Finding the trending venues in each Neighborhood

Foursquare location data is used to retrieve top 5 common venues in each neighborhood. First, the data frame was modified to add the *categories* column. Then, using the Foursquare API each neighborhood was explored and the venues retrieved. The venues of each neighborhood were within a 500-meter radius of the neighborhood's coordinate. A limit of 100 venues for each neighborhood was set. The retrieved values were put into a data frame as shown in figure 7.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Adams-Normandie	34.033081	-118.297115	Ignatius Cafe	34.031772	-118.293006	Café
1	Adams-Normandie	34.033081	-118.297115	Orange Door Sushi	34.032485	-118.299368	Sushi Restaurant
2	Adams-Normandie	34.033081	-118.297115	Shell	34.033095	-118.300025	Gas Station
3	Adams-Normandie	34.033081	-118.297115	Sushi Delight	34.032445	-118.299525	Sushi Restaurant
4	Adams-Normandie	34.033081	-118.297115	Little Xian	34.032292	-118.299465	Sushi Restaurant

Figure 7. Top 5 common venues for each neighborhood.

The frequency for each venue needed to be calculated next. The venue with the highest frequency is considered the most common venue. To calculate the frequency, the data frame above was used to create a one-hot array and grouped by the neighborhoods to produce the table below.

```
----Adams-Normandie----
              venue  freq
0  Sushi Restaurant  0.33
1              Food  0.11
2       Taco Place  0.11
3      Gas Station  0.11
4             Café  0.11


----Agoura Hills----
              venue  freq
0     Deli / Bodega  0.12
1  Indian Restaurant  0.12
2    Breakfast Spot  0.12
3       Gas Station  0.12
4               Gym  0.12
```

Figure 8. The top 5 venues and frequency for each neighborhood.

After calculating the frequency of all the venues in each neighborhood, the top 4 common venues table in figure 8 was converted into a data frame shown in figure 9.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
0	Adams-Normandie	Sushi Restaurant	Café	Taco Place	Donut Shop
1	Agoura Hills	Gym	Pharmacy	Deli / Bodega	Gas Station
2	Agua Dulce	Grocery Store	Pizza Place	Bakery	Café
3	Alhambra	Ice Cream Shop	Bakery	Burger Joint	Café
4	Alondra Park	Gym / Fitness Center	Park	Baseball Field	Fabric Shop

Figure 9. Data frame containing neighborhoods and their 5 most common venues.

Step 5 - Selecting Neighborhoods which has Coffee Shops as the most visited place

Selecting the neighborhoods which have ‘coffee shop’ as their 1st most common venue from the data frame in figure 9. The new data frame has 13 rows and 5 columns as shown in the figure below.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
0	Atwater Village	Coffee Shop	Theater	Pizza Place	Pet Store
1	Century City	Coffee Shop	Food Truck	Italian Restaurant	Salad Place
2	Claremont	Coffee Shop	Bakery	Ice Cream Shop	American Restaurant
3	Culver City	Coffee Shop	Italian Restaurant	Gastropub	American Restaurant
4	Larchmont	Coffee Shop	Mexican Restaurant	Spa	Café
5	Los Feliz	Coffee Shop	Pizza Place	Kids Store	American Restaurant
6	Monrovia	Coffee Shop	Mexican Restaurant	Japanese Restaurant	Spa
7	Norwalk	Coffee Shop	Ice Cream Shop	Department Store	Bank
8	Pacific Palisades	Coffee Shop	Ice Cream Shop	American Restaurant	Bank
9	Santa Monica	Coffee Shop	Café	Mexican Restaurant	Shipping Store
10	Sawtelle	Coffee Shop	Yoga Studio	Sandwich Place	Movie Theater
11	South Park	Coffee Shop	American Restaurant	Sports Bar	Food Truck
12	Windsor Square	Coffee Shop	Italian Restaurant	Juice Bar	Sushi Restaurant

Figure 10. New data frame which contains the neighborhoods with ‘coffee shop’ as 1st most common venue.

Step 6 - Finding and adding the population in each Neighborhood

The population of the 13 selected neighborhoods is read from an excel file into a data frame. Then it is merged with the data frame in Figure 10 using inner join to produce the data frame below.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	Population
0	Atwater Village	Coffee Shop	Pizza Place	Pet Store	Restaurant	29092
1	Century City	Coffee Shop	Food Truck	Italian Restaurant	Salad Place	8554
2	Claremont	Coffee Shop	American Restaurant	Bakery	Arts & Crafts Store	36015
3	Culver City	Coffee Shop	Gastropub	American Restaurant	Thai Restaurant	39283
4	Larchmont	Coffee Shop	Mexican Restaurant	Movie Theater	Café	62733

Figure 11. Neighborhoods, their common venues, and population

To visualize the population of each neighborhood, a bar chart was plotted.

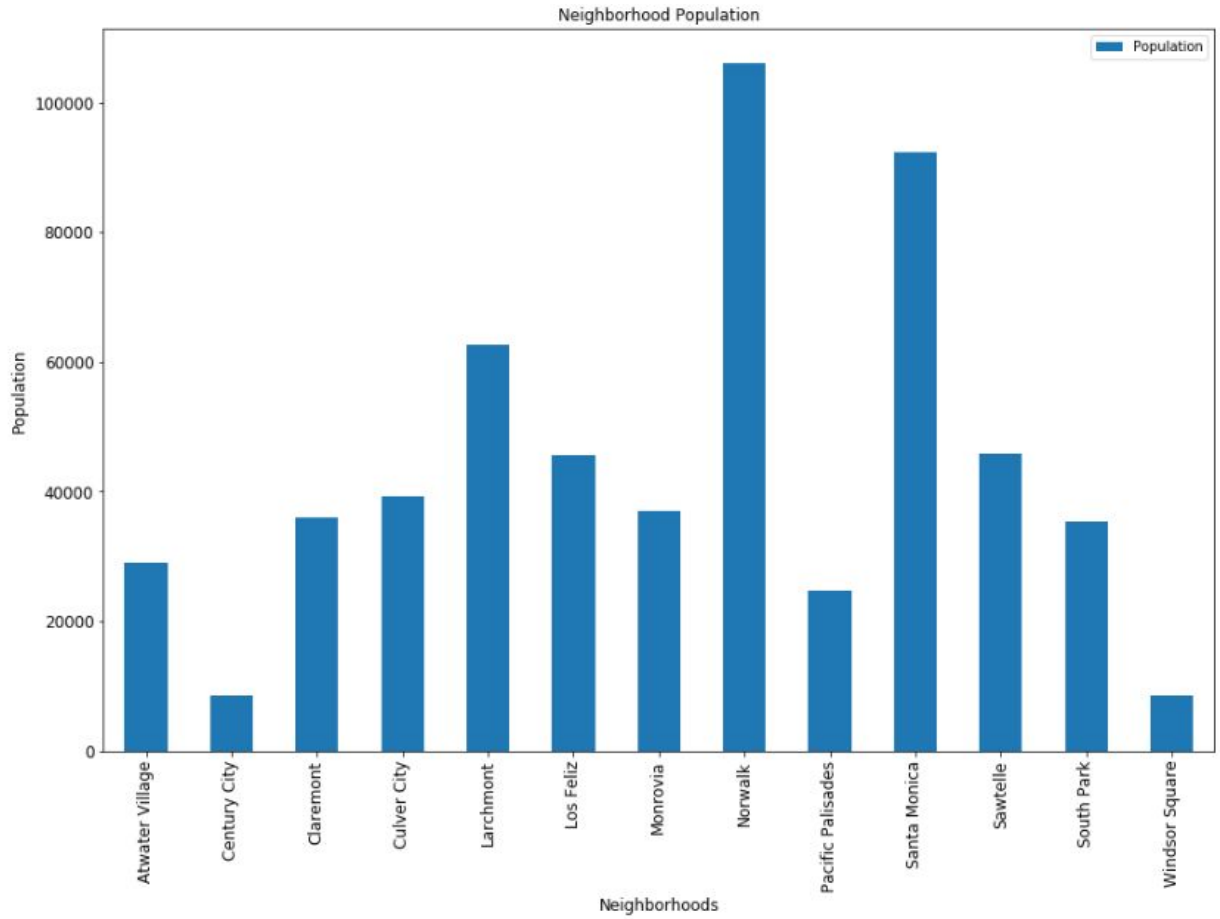

Figure 12. Neighborhood population.

Step 7 - Shortlisting to 2 Neighborhoods with the highest population

From the 13 neighborhoods, the top 2 neighborhoods with the highest population were selected.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	Population
7	Norwalk	Coffee Shop	Ice Cream Shop	Department Store	Bank	106084
9	Santa Monica	Coffee Shop	Café	Mexican Restaurant	Shipping Store	92306

Figure 13. The 2 neighborhoods with the highest population.

Step 8 - Finding the number of Starbucks in the neighborhoods

To identify the number of Starbucks' in Norwalk and Santa Monica, Foursquare location data was used again. All the Starbucks within 10000 meter-radius of Norwalk's coordinate was located. Then,

only the Starbucks' which are in the neighborhood, Norwalk, were selected and counted. The same methods were used to select and count the number of Starbucks in the Santa Monica neighborhood.

Step 9 - The Final neighborhood

The map of Santa Monica was visualized using the folium library. Again, using the Foursquare location data all the Starbucks located in Santa Monica were marked on the map. Then the locations of all the Starbucks' were put into two clusters and the number of Starbucks' in each cluster was determined.

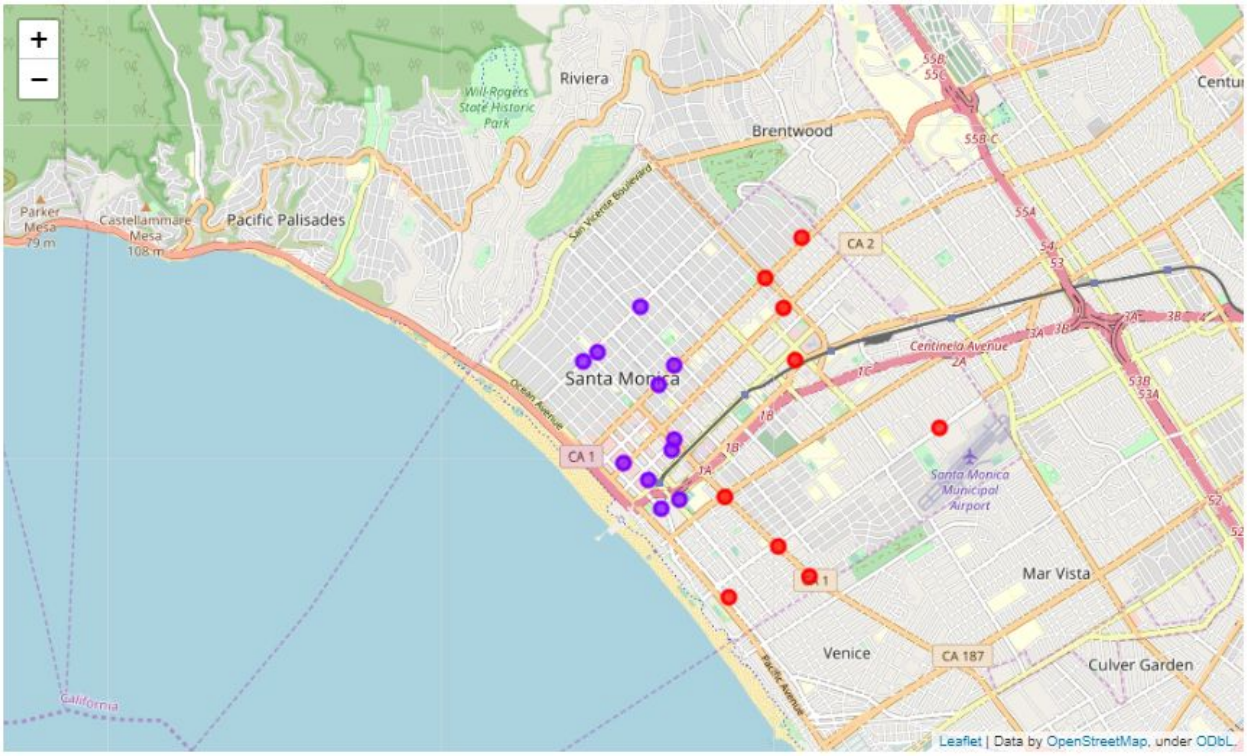

Figure 13. Starbucks' location in Santa Monica put into 2 clusters.

Results

There are, in total 272, neighborhoods in the Los Angeles county. Out of 272, only 13 neighborhoods have their 1st most common venue as coffee shops.

	Neighborhood	1st Most Common Venue
0	Atwater Village	Coffee Shop
1	Century City	Coffee Shop
2	Claremont	Coffee Shop
3	Culver City	Coffee Shop
4	Larchmont	Coffee Shop
5	Los Feliz	Coffee Shop
6	Monrovia	Coffee Shop
7	Norwalk	Coffee Shop
8	Pacific Palisades	Coffee Shop
9	Santa Monica	Coffee Shop
10	Sawtelle	Coffee Shop
11	South Park	Coffee Shop
12	Windsor Square	Coffee Shop

Figure 14. The 13 neighborhoods with 1st most common venue ‘coffee shop’.

Among these 13 neighborhoods, Norwalk and Santa Monica have the highest number of people living there. Norwalk has a population of 106,084 residents and Santa Monica has a population of 92,306 residents.

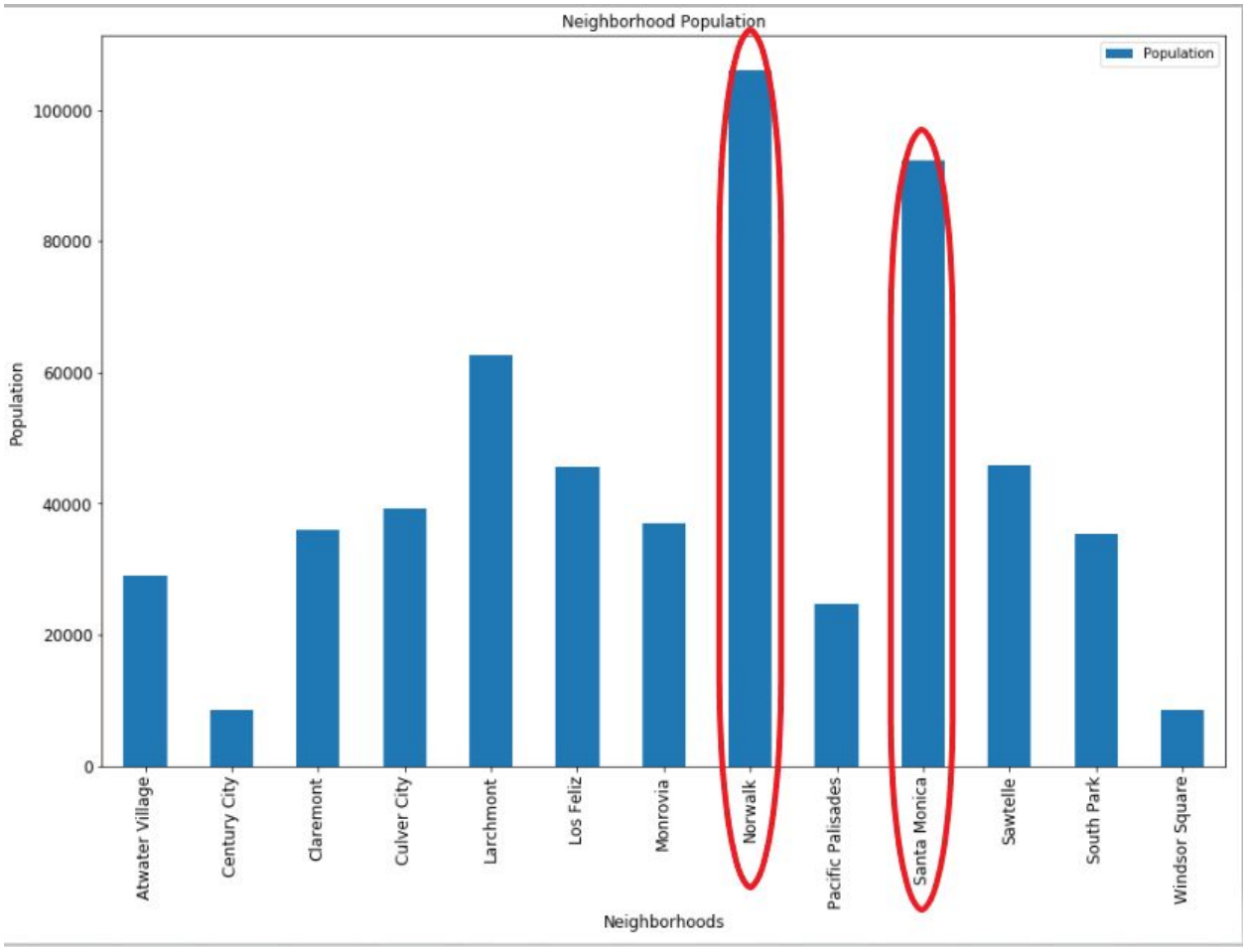

Figure 15. Norwalk and Santa Monica are the 2 neighborhoods with the highest population.

Among those two neighborhoods, Santa Monica has a higher number of Starbucks with 20. But Norwalk has a higher number of customers per store, 11,787 customers per store.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	Population	Starbucks	Customers per Shop
6	Norwalk	Coffee Shop	Multiplex	Ice Cream Shop	Department Store	106084	9	11787
7	Santa Monica	Coffee Shop	Café	Bank	Mexican Restaurant	92306	20	4615

Figure 16. Number of Starbucks in Norwalk and Santa Monica.

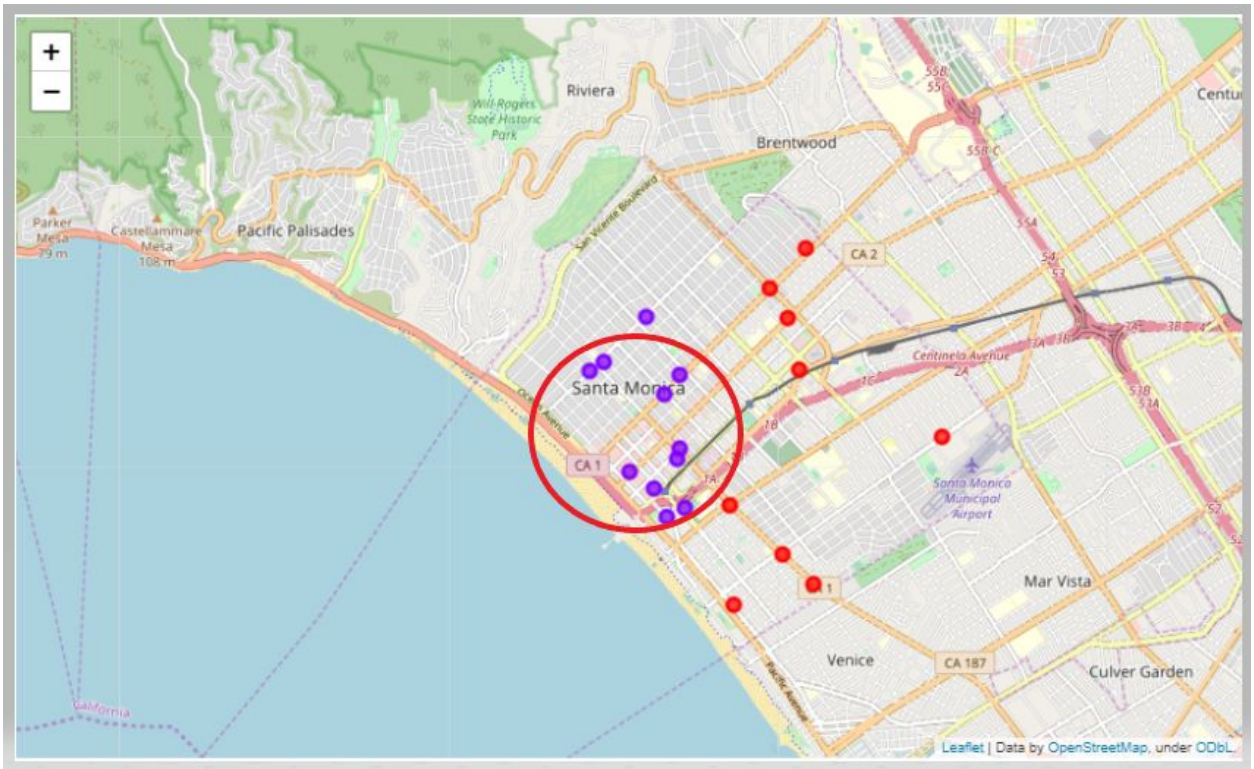

Figure 17. Red circle showing the where there is a higher concentration of Starbucks'.

There are 11 Starbucks in the purple cluster and 9 in the red cluster.

Discussion

The intention of this project was to find the best neighborhood, in Los Angeles, and an area within that neighborhood, ideal for Tim Hortons to open its first shop. So out of 272 neighborhoods, it needed to be shortlisted to 1.

With the top 5 most common venues of the 272 neighborhoods information retrieved, it was possible to shortlist to 13 neighborhoods. These 13 neighborhoods have their 1st most common venue as 'coffee shop'. Meaning coffee shops are very popular in these neighborhoods and there will be a good customer base.

Then, the population of all 13 neighborhoods was taken into account to further narrow down the list. Visualizing this information, figure 15, it was clear that two neighborhoods had a much higher population than the others. So now it was down to two neighborhoods, Norwalk and Santa Monica. A higher population of the neighborhood means more potential customer for the shop.
The last method used to choose between the 2 neighborhoods was by counting the number of Starbucks in each of the neighborhood. The higher number of Starbucks means more sales for coffee shops in the neighborhood. The more sales the company makes, the more they expand. Starbucks is one of Tim Hortons' biggest competition. If there are more sales for Starbucks in one

neighborhood, there should be more sales for Tim Hortons in that neighborhood. So Santa Monica, with a higher number of Starbucks, is the final state.

The last step was to specify the area in the neighborhood where the shop should be located. To do this, the location of Starbucks' in Santa Monica was put into 2 clusters. The cluster with more Starbucks' means that the area has more potential customers. Therefore, that is the cluster in which Tim Hortons should have their shop in. The purple cluster in figure 17.

Conclusion

The purpose of this project was to find a location in Los Angeles county where Tim Hortons should open their shop. Starting with 272 neighborhoods, it was narrowed down into 1. The selection was done by using the information on popular venues, population and the amount of Starbucks' in neighborhoods. The final neighborhood was Santa Monica. Santa Monica's most popular venues are coffee shops, with Starbucks having 20 shops. Santa Monica also has a population of 92,306. Finally, with clustering, the area inside the neighborhood where Tim Hortons should have a shop was determined. This area, marked with a red circle in figure 17, has a higher density of coffee shops.

In conclusion, if Tim Hortons opens a shop in Los Angeles, that area in Santa Monica will be the most ideal location.

References

[1]"Tim Hortons takes on the world | The Star", *thestar.com*, 2019. [Online]. Available: https://www.thestar.com/business/2019/05/15/tim-hortons-takes-on-the-world.html#:~:targetText=Tim%20Hortons%2C%20Canada's%20largest%20quick,outlets%20in%20Canada%20in%202017. [Accessed: 26- Nov- 2019].

[2]"Tim Hortons pulls back on U.S. expansion amid franchisee fracas - BNN Bloomberg", *BNN*, 2019. [Online]. Available: https://www.bnnbloomberg.ca/tim-hortons-pulls-back-on-u-s-expansion-amid-franchisee-fracas-1.1087178. [Accessed: 26- Nov- 2019].

[3]M. Tips and T. Population, "The Top 10 Largest U.S. Cities by Population | Moving.com", *Moving.com*, 2019. [Online]. Available: https://www.moving.com/tips/the-top-10-largest-us-cities-by-population/. [Accessed: 26- Nov- 2019].

[4]2019. [Online]. Available: https://craft.co/tim-hortons/competitors#:~:targetText=Tim%20Hortons's%20top%20competitors%20include,%2C%20Dunkin'%20Brands%20and%20Starbucks.&targetText=Tim%20Hortons%20is%20a%20multinational,for%20its%20coffee%20and%20donuts.&targetText=Burger%20King%20is%20a%20global%20chain%20

of%20hamburger%20fast%20food%20restaurants.&targetText=McDonald's%20is%20an%20international%20chain%20of%20hamburger%20fast%20food%20restaurants. [Accessed: 26- Nov- 2019].