

The Battle of Neighborhood

IBM Capstone Project

Javari Food Company LLC



Introduction

- Javari Food Company LLC looking to open new Mediterranean restaurant in Canada
- Toronto is chosen to open the first restaurant in Canada

Business Problem:

- Location critical for company's success in Canada
- Easy to replicate
- Low Competition

Success Criteria:

- Goal is to identify optimal locations for the new restaurant

Toronto – City Facts

- Most populous city in Canada
- It is diverse and financial capital of Canada
- It is diverse in ethnicity and culture
- Business friendly environment
- Highly competitive
- High cost of doing business
- High diverse economy with strengths in multiple areas such as technology, financial services, banking etc.



Toronto – Food Industry

- City is famous for excellent cuisines.
- Food culture includes an array of international cuisines influenced by the city's immigrant history.
- Food industry is highly competitive.
- Factor that influence success of a restaurant in this city:
 - Location
 - Competition
 - Demographics
 - Menu



Factors to consider for decision making

- City Population
- City Demographics
- Nearby entertainment and attractions
- Competitors in that location
- Type of Cuisine and menu served
- Market is saturated or untapped

Data – 1 – Neighborhood Information

- Toronto Postal codes and neighborhood information retrieved from the following link:
https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

| | Postcode | Borough | Neighbourhood |
|---|----------|-------------|---------------------------------------|
| 0 | M1B | Scarborough | Rouge,Malvern |
| 1 | M1C | Scarborough | Highland Creek ,Rouge Hill,Port Union |
| 2 | M1E | Scarborough | Guildwood,Morningside,West Hill |
| 3 | M1G | Scarborough | Woburn |
| 4 | M1H | Scarborough | Cedarbrae |

Data – 2 – Geospatial coordinates

- Toronto city geospatial coordinates data will be utilized as input for the Foursquare API
- The geospatial information is retrieved from the link:
http://cocl.us/Geospatial_data

| | Postcode | Latitude | Longitude |
|---|----------|-----------|------------|
| 0 | M1B | 43.806686 | -79.194353 |
| 1 | M1C | 43.784535 | -79.160497 |
| 2 | M1E | 43.763573 | -79.188711 |
| 3 | M1G | 43.770992 | -79.216917 |
| 4 | M1H | 43.773136 | -79.239476 |

Data – 3 – Population and Demographics

- Toronto population from the census data gathered every 5 years from 1996 to 2016 taken from the following link:

https://en.wikipedia.org/wiki/Demographics_of_Toronto

| | 2016 | 2011 | 2006 | 2001 | 1996 |
|----------------|--------|----------|----------|----------|----------|
| Ethnic Origins | | | | | |
| Canadian | 746960 | 728745.0 | 651635.0 | 861945.0 | 710755.0 |
| English | 732555 | 777110.0 | 804100.0 | 783770.0 | 891735.0 |
| Chinese | 700705 | 594735.0 | 537060.0 | 435685.0 | 359450.0 |
| Indian | 643370 | 572250.0 | 484655.0 | 345855.0 | 255685.0 |
| Irish | 544380 | 543600.0 | 531865.0 | 487215.0 | 480980.0 |
| Scottish | 543760 | 545365.0 | 561050.0 | 517115.0 | 534595.0 |
| Italian | 484360 | 475090.0 | 466155.0 | 429385.0 | 414310.0 |
| Filipino | 274675 | 246345.0 | 181330.0 | 140405.0 | 102525.0 |
| German | 271815 | 262830.0 | 259015.0 | 220140.0 | 224525.0 |
| French | 247790 | 249375.0 | 241395.0 | 220535.0 | 236315.0 |

Data – 4 – Venue information -- Foursquare.com

- City geospatial coordinates data will be utilized as input for the Foursquare API, that will be leveraged to provision venues information for each neighborhood.
- Foursquare API will be utilized to explore neighborhoods in the City.

| | Neighbourhood | Neighbourhood Latitude | Neighbourhood Longitude | Venue | Venue Latitude | Venue Longitude | Venue Category |
|---|---------------|------------------------|-------------------------|------------------------------------|----------------|-----------------|-------------------|
| 0 | The Beaches | 43.676357 | -79.293031 | The Big Carrot Natural Food Market | 43.678879 | -79.297734 | Health Food Store |
| 1 | The Beaches | 43.676357 | -79.293031 | Grover Pub and Grub | 43.679181 | -79.297215 | Pub |
| 2 | The Beaches | 43.676357 | -79.293031 | Starbucks | 43.678798 | -79.298045 | Coffee Shop |
| 3 | The Beaches | 43.676357 | -79.293031 | Glen Stewart Park | 43.675278 | -79.294647 | Park |
| 4 | The Beaches | 43.676357 | -79.293031 | Upper Beaches | 43.680563 | -79.292869 | Neighborhood |

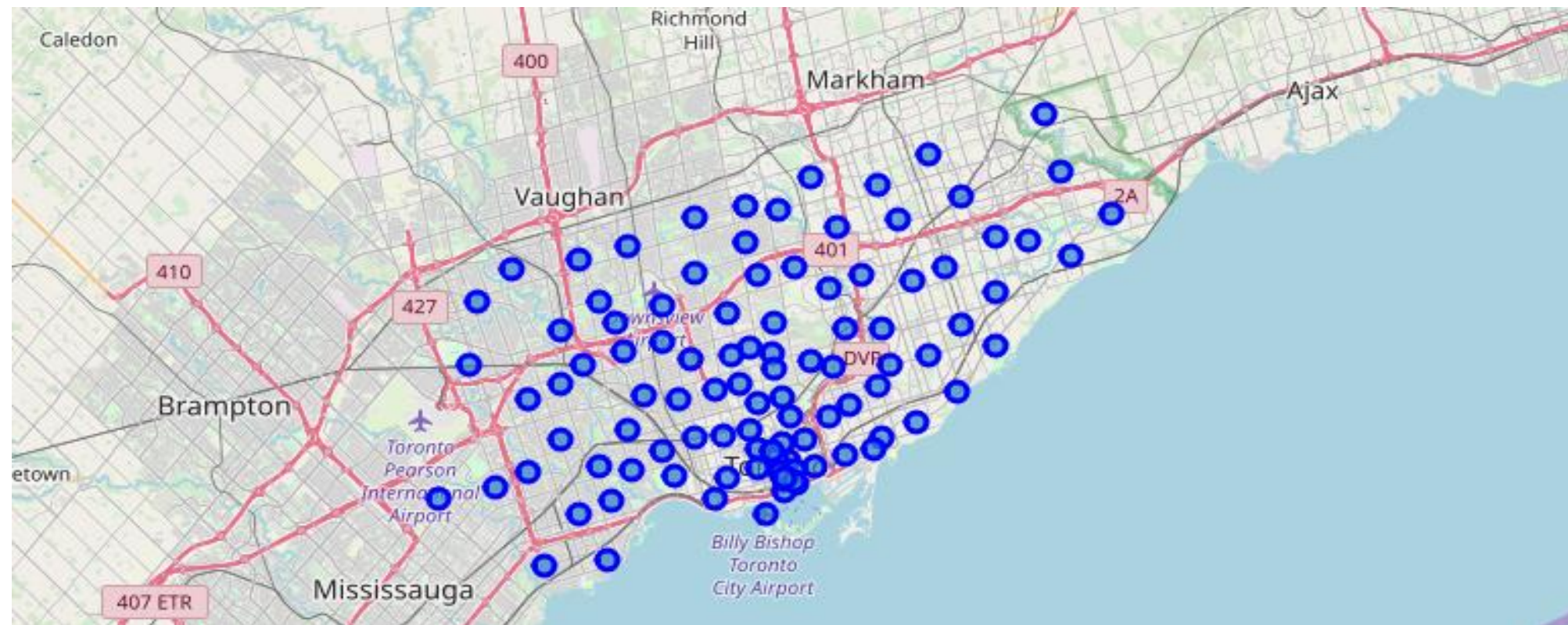
Analytical Approach

- Neighborhoods from 4 Boroughs in Toronto chosen for study:
 - Downtown Toronto
 - East Toronto
 - West Toronto
 - Central Toronto
- Venue data retrieved for these neighborhoods using Foursquare API
- Only restaurant data is filtered from the venues data and utilized for analysis for this project.

Exploratory analysis – Data-1

Toronto City – Geospatial coordinates data

- Toronto's neighborhood data is merged with Geospatial data using Pandas dataframe
- Data will be used to get venues information using Foursquare API
- Folium map libraries used to create map of Toronto city with neighborhoods superimposed on top.

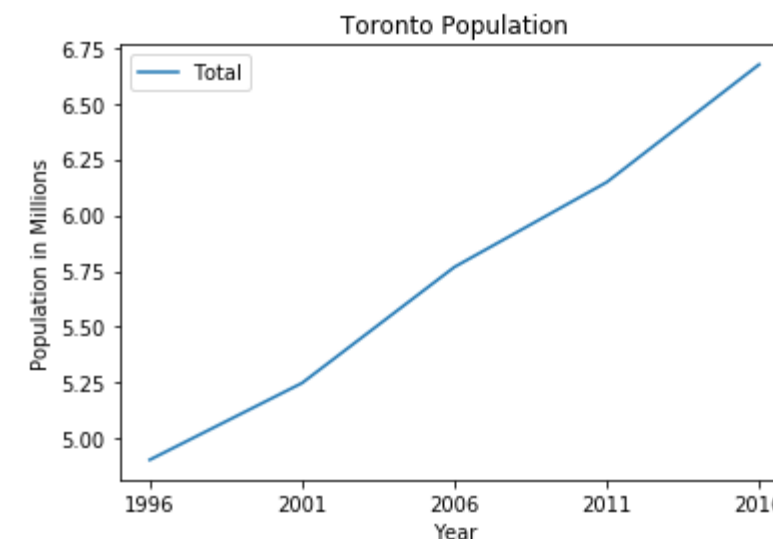


Exploratory analysis – Data-2

Toronto City – Population and Demographics

- Population census data is retrieved from the Wikipedia link and loaded to Pandas dataframe
- Data used to analyze different ethnic origins living in the city over the period
- Population in the city is steadily increasing which indicates the more demand for restaurants.

| | 2016 | 2011 | 2006 | 2001 | 1996 |
|----------------|--------|----------|----------|----------|----------|
| Ethnic Origins | | | | | |
| Canadian | 746960 | 728745.0 | 651635.0 | 861945.0 | 710755.0 |
| English | 732555 | 777110.0 | 804100.0 | 783770.0 | 891735.0 |
| Chinese | 700705 | 594735.0 | 537060.0 | 435685.0 | 359450.0 |
| Indian | 643370 | 572250.0 | 484655.0 | 345855.0 | 255685.0 |
| Irish | 544380 | 543600.0 | 531865.0 | 487215.0 | 480980.0 |



Exploratory analysis – Data-3

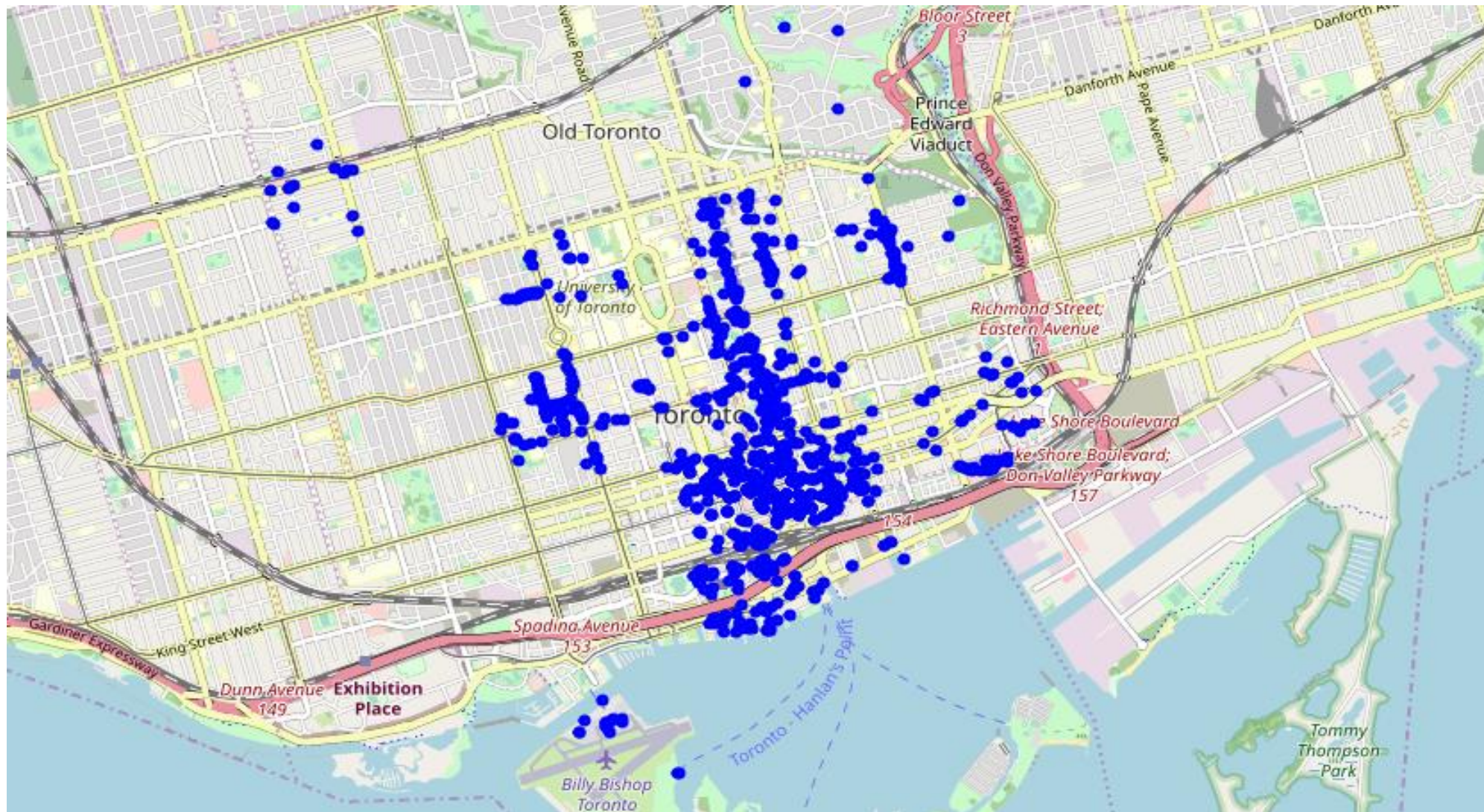
Venues data using Foursquare API

- Toronto neighborhood and geospatial data used as an input to the Foursquare API
- Used venues data returned by the API to explore neighborhoods in the city
- Venues data from neighborhoods in 4 boroughs:

| | Neighbourhood | Neighbourhood Latitude | Neighbourhood Longitude | Venue | Venue Latitude | Venue Longitude | Venue Category |
|---|---------------|------------------------|-------------------------|------------------------------------|----------------|-----------------|-------------------|
| 0 | The Beaches | 43.676357 | -79.293031 | The Big Carrot Natural Food Market | 43.678879 | -79.297734 | Health Food Store |
| 1 | The Beaches | 43.676357 | -79.293031 | Grover Pub and Grub | 43.679181 | -79.297215 | Pub |
| 2 | The Beaches | 43.676357 | -79.293031 | Starbucks | 43.678798 | -79.298045 | Coffee Shop |
| 3 | The Beaches | 43.676357 | -79.293031 | Glen Stewart Park | 43.675278 | -79.294647 | Park |
| 4 | The Beaches | 43.676357 | -79.293031 | Upper Beaches | 43.680563 | -79.292869 | Neighborhood |

Venues Visualization of Downtown Toronto

- Downtown Toronto consists of 749 venues and 204 unique venue types



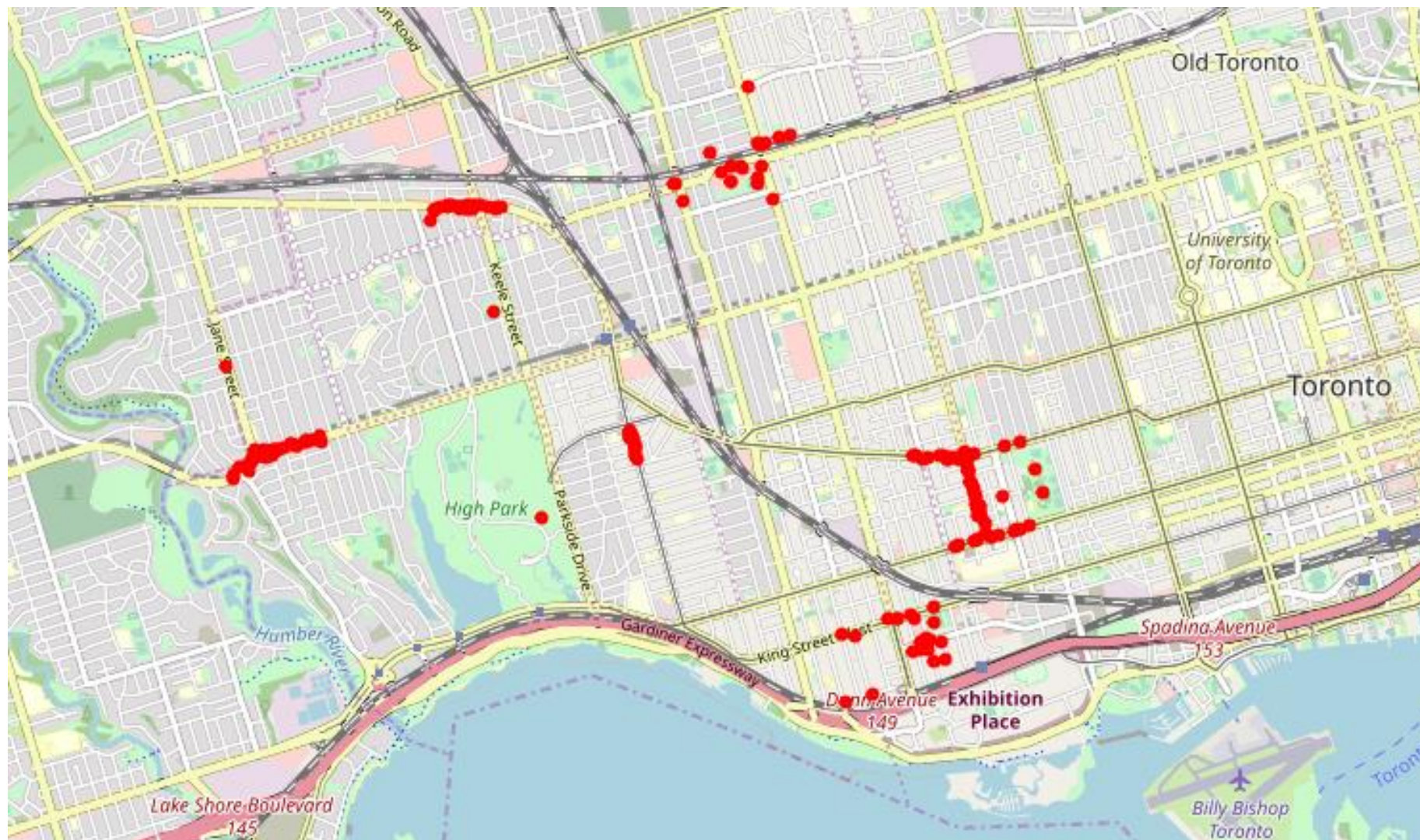
Venues Visualization of East Toronto

- East Toronto consists of 121 venues and 68 unique venue types



Venues Visualization of West Toronto

- West Toronto consists of 176 venues and 90 unique venue types



Venues Visualization of Central Toronto

- Central Toronto consists of 102 venues and 64 unique venue types



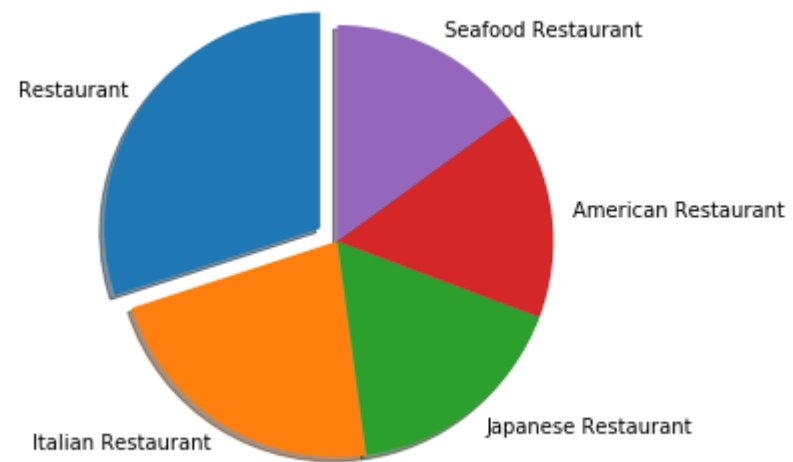
Restaurants analysis in Toronto

- Restaurants were filtered from the venue data pulled from Foursquare API for analysis.
- Extracted restaurant type for each neighborhood and grouped by borough
- Identified Top 5 restaurants in each borough to give management an idea on the popularity in each area.

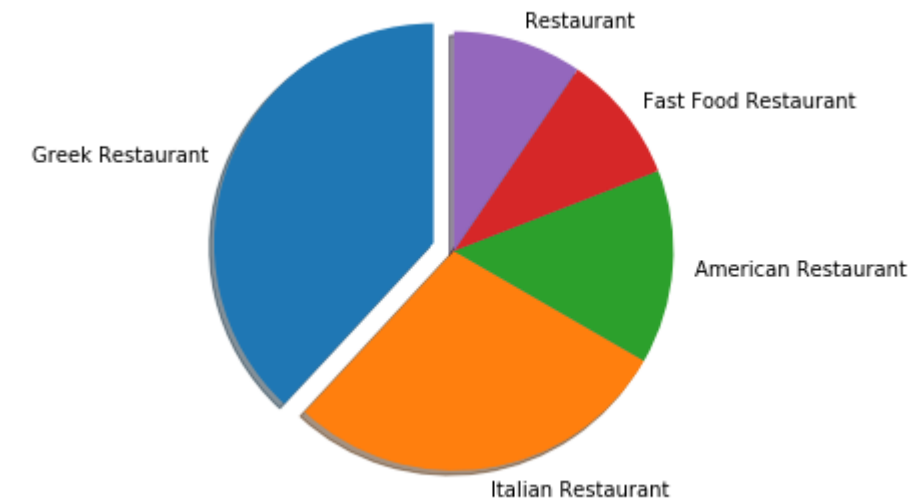


Top 5 restaurants in each Borough in Toronto

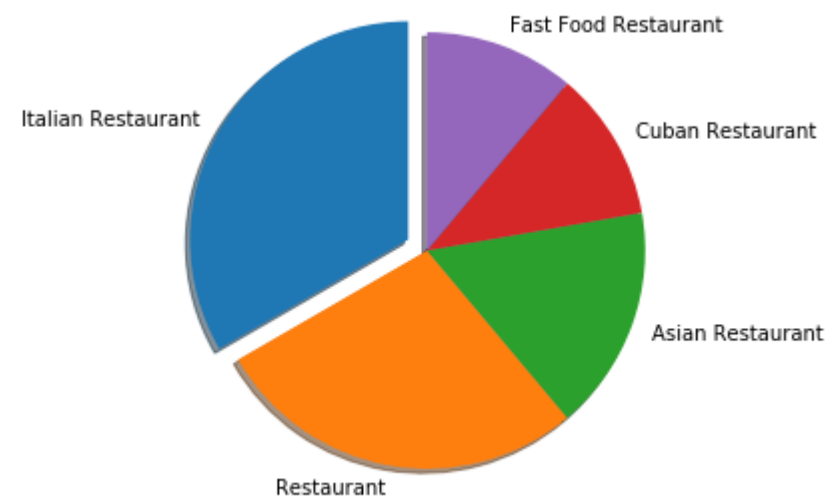
Top 5 restaurants in Downtown Toronto



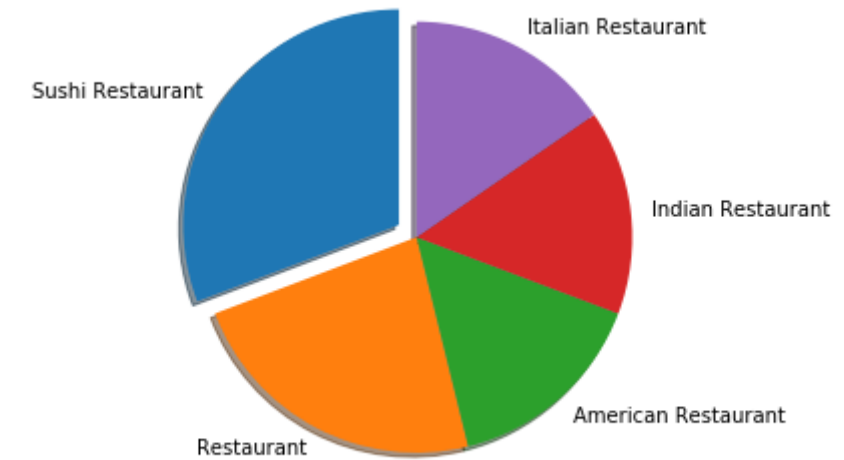
Top 5 restaurants in East Toronto



Top 5 restaurants in West Toronto



Top 5 restaurants in Central Toronto

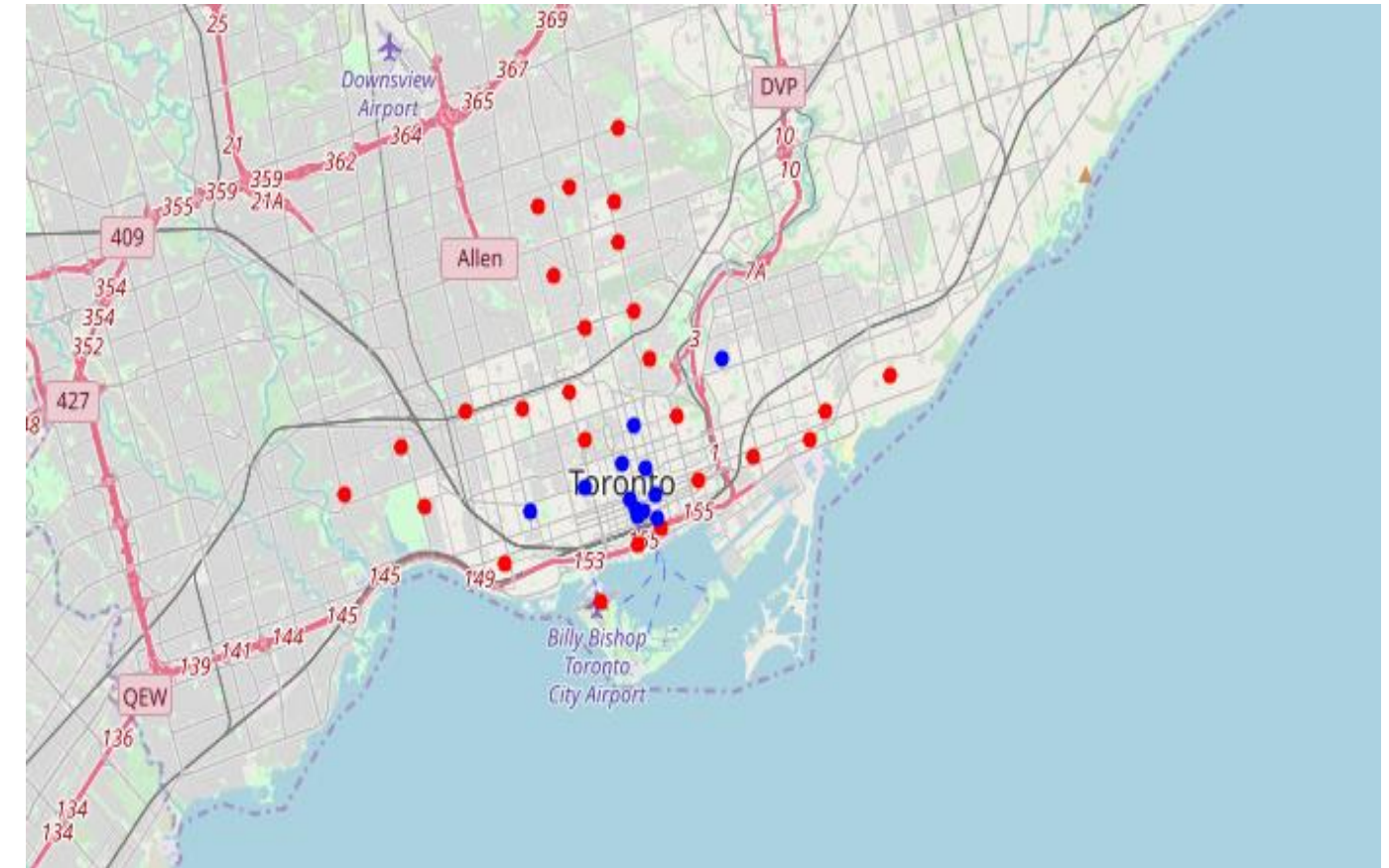


Neighborhood *K*-means clustering

- K-means algorithm used to cluster the neighborhoods
- Algorithm partitions n observations into K clusters
- Each observation belongs to the cluster with the nearest mean
- Silhouette coefficient method used to determine the optima value of K
- Identified that Silhouette coefficient was higher when $K=2$

Results

- Used $K=2$ and ran the K-means clustering algorithm
- **Cluster 0:**
 - Total and Total Sum of Cluster0 has highest value.
 - Markets are saturated. Number of restaurants are very high in this cluster.
- **Cluster 1:**
 - Total and Total Sum of Cluster1 has smallest value.
 - Markets are not saturated. There are some neighborhoods with no restaurants.



| | Total | Total Sum |
|----------|-----------|-----------|
| cluster0 | 24.500000 | 49.000000 |
| cluster1 | 4.461538 | 8.923077 |

Cluster 1 – Neighborhoods with no restaurants

East, Central and Downtown Toronto all have some neighborhoods that has no restaurants.

| | Borough | Neighbourhood | Latitude | Longitude | Total | Cluster_Labels |
|---|------------------|---|-----------|------------|-------|----------------|
| 0 | East Toronto | The Beaches | 43.676357 | -79.293031 | 0 | 1 |
| 1 | Central Toronto | Lawrence Park | 43.728020 | -79.388790 | 0 | 1 |
| 2 | Central Toronto | Davisville North | 43.712751 | -79.390197 | 0 | 1 |
| 3 | Downtown Toronto | Rosedale | 43.679563 | -79.377529 | 0 | 1 |
| 4 | Central Toronto | Roselawn | 43.711695 | -79.416936 | 0 | 1 |
| 5 | Downtown Toronto | CN Tower,Bathurst Quay,Island airport,Harbourf... | 43.628947 | -79.394420 | 0 | 1 |

Discussion

- Toronto is a big city with a high population density
- Very diverse ethnic and cultural background
- Focused on neighborhoods in 4 boroughs – Downtown, East, West and Central Toronto.
- used K-means algorithm as part of this clustering study and identified the optimum K value as 2
- Identified neighborhoods with no restaurants which shows that there is no nearby competition
- Locations can be used as markers/starting points for more detailed local analysis based on other factors

Conclusion

- Goal is to identify neighborhoods in the City of Toronto with less or no restaurants.
- Identified all venues in each neighborhood in the city, filtered the restaurants.
- Clustering was then performed to create zones of interest and later identified the neighborhoods that has no restaurants.
- Final decision on optimal restaurant location will be made by management based on specific characteristics of neighborhoods.
- Analysis provided will aid the decision making process to identify a venue with lower risk and competition.

