Capstone Project Presentation

Project: Exploring Toronto for Opening a New Chain of Gourmet Pizza Place with Bar

Shahbaz Masih

Summary

- **Introduction**
  - Business Problem
  - Target Audience

- **Data**
  - Required Data
  - Data Sources

- **Methodology**

- **Results**

- **Discussion**

- **Conclusion**

Introduction

- **Business Problem**
  - What are the best possible locations to open a chain of gourmet pizza places with the facility of a bar.

- **Target Audience**
  - The entrepreneurs who want to find the most suitable location to open a pizza place with a bar.

Data

- **Required Data**
  - A List of neighborhoods in Toronto
  - Geographical locations (Latitudes and Longitudes) of these neighborhoods
  - The distribution of population by different ethnicities
  - Data of the venues (pizza places and bars) in Toronto
- **Data Sources**
  - "https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)" Wiki page
  - "https://cocl.us/Geospatial_data"
  - "Demographics of Toronto" (https://en.m.wikipedia.org/wiki/Demographics_of_Toronto#Ethnic_diversity)
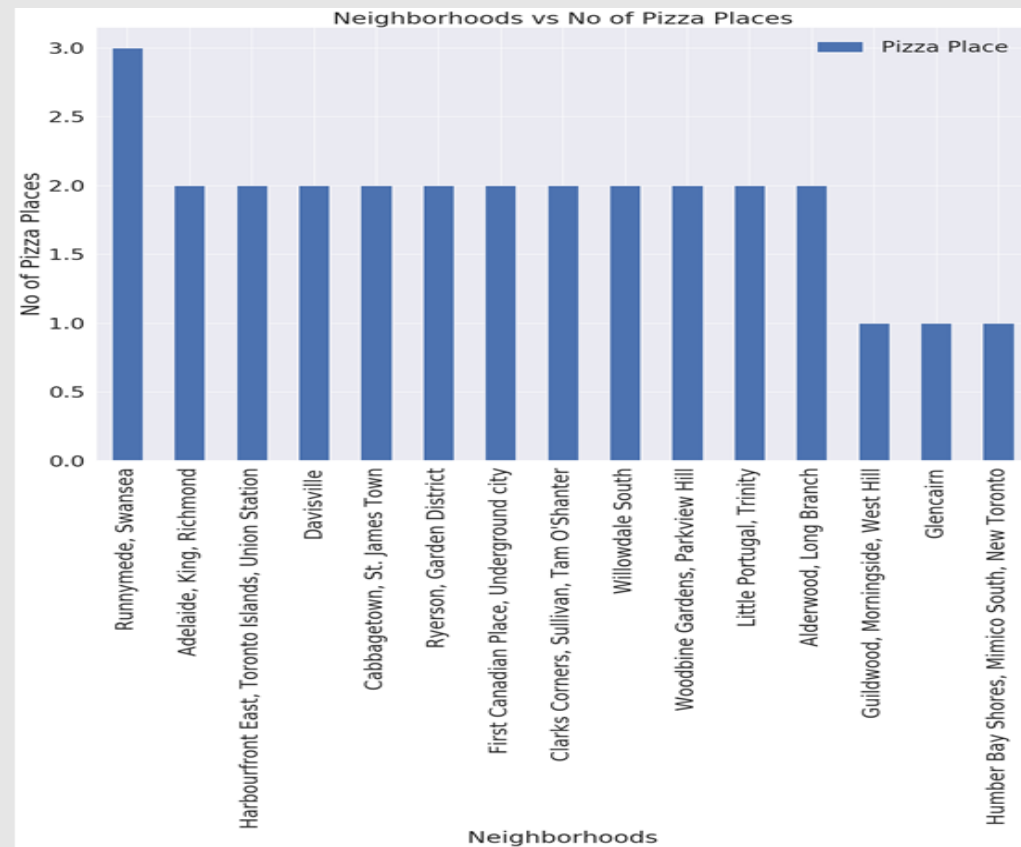 - Foursquare's API ((https://developer.foursquare.com/docs)

Methodology

- **Data Wrangling**
  - Data acquisition by using online sources and provided list of postal to find venues in each neighbourhood and Italian population distribution
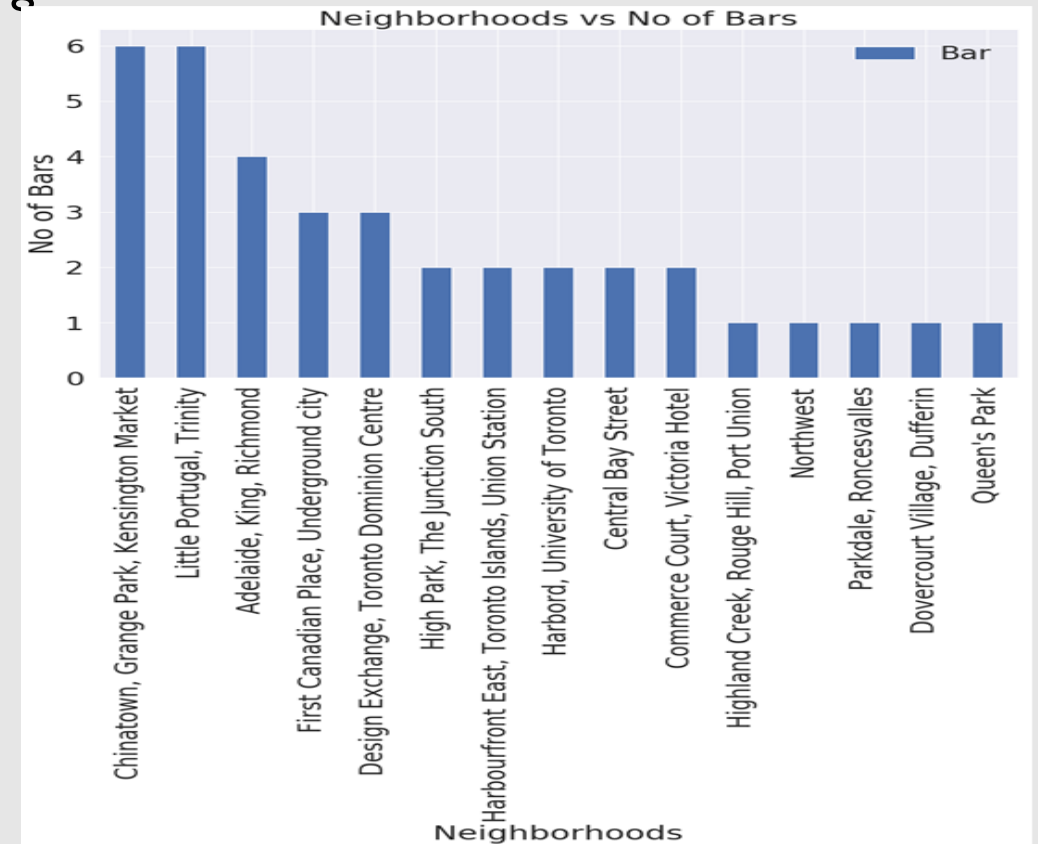
Methodology
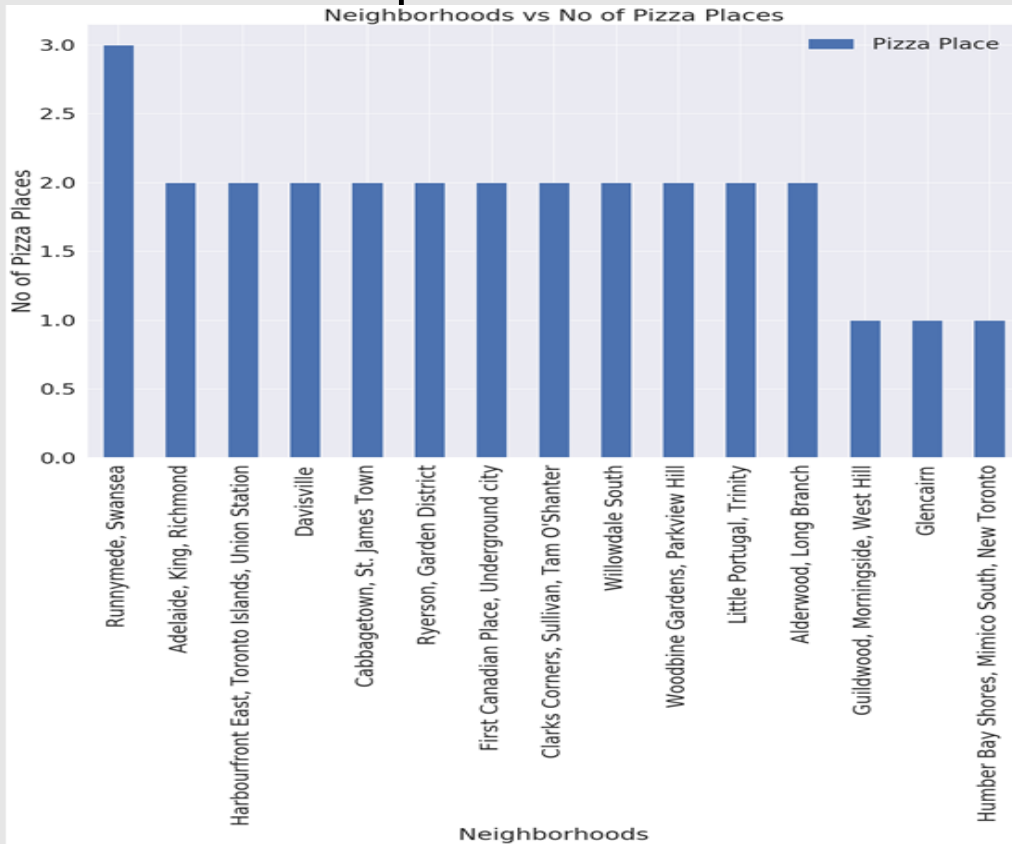
- **Exploratory Data Analysis**
  - Number of Pizza places in different neighbourhoods
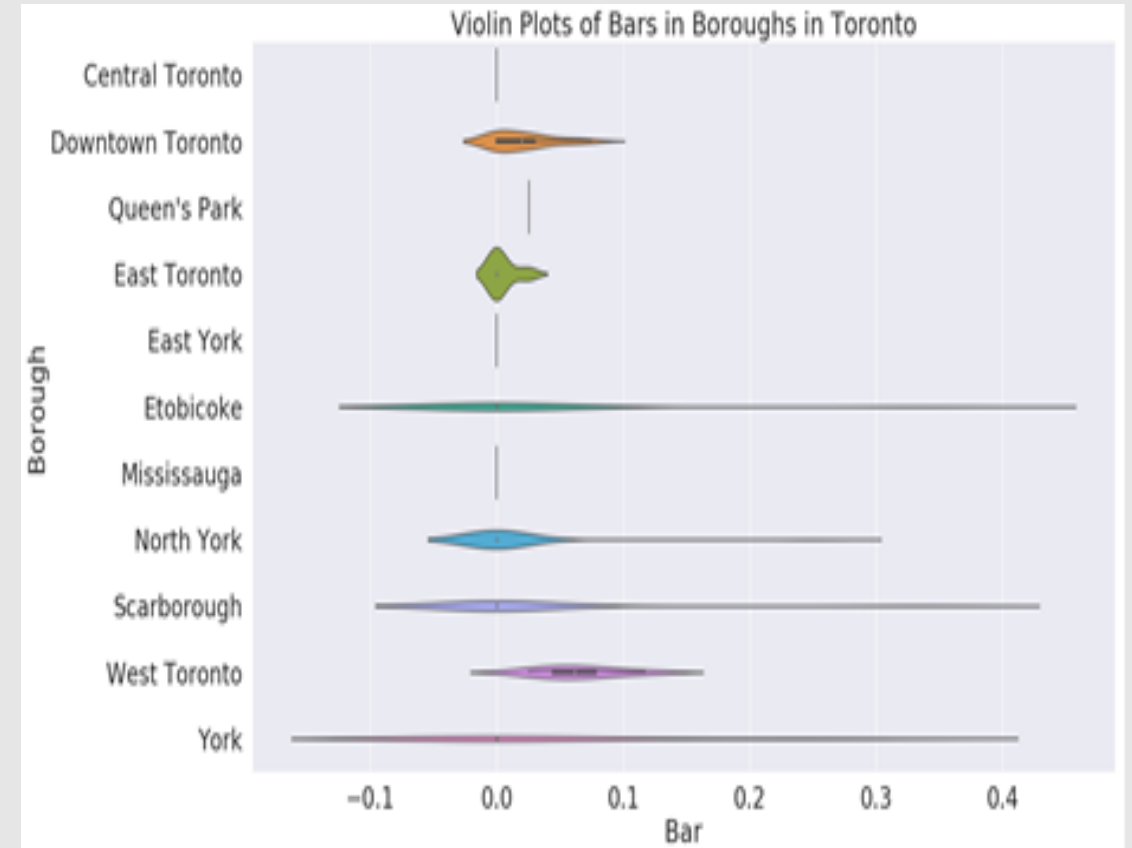

Neighborhoods vs No of Pizza Places

Methodology

◦ **Exploratory Data Analysis**

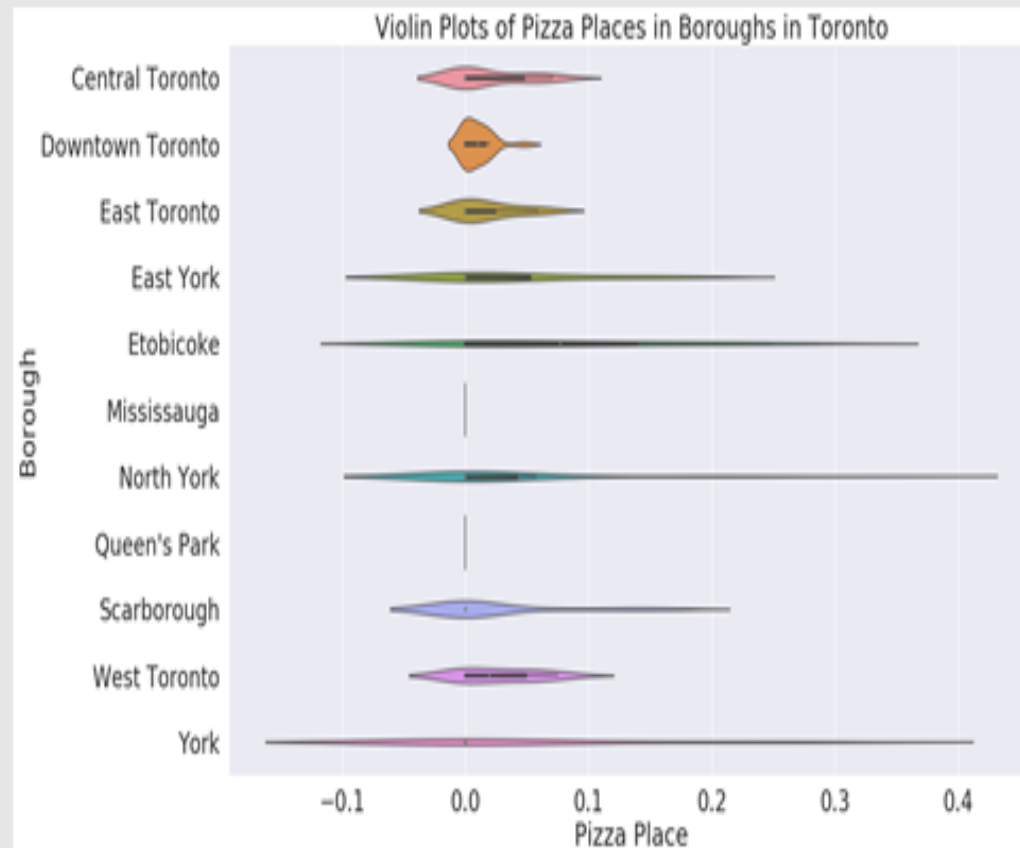   ◦ Number of Pizza places and bars in different neighbourhoods

Methodology

◦ **Exploratory Data Analysis**

  ◦ Number of Pizza places and distribution in different neighbourhoods

Methodology

○ **Exploratory Data Analysis**

    ○ Italian population distribution in different neighbourhoods
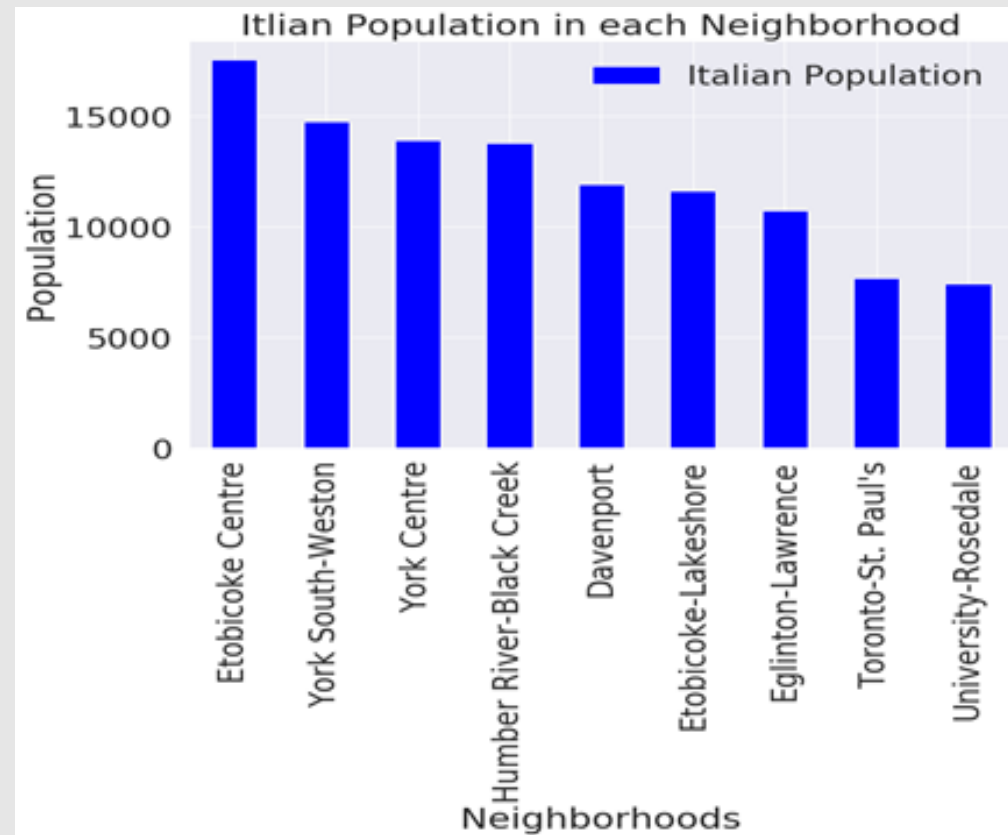
Methodology

- **Use of Machine Learning and choice of K-Means Clustering**
  - Determining optimum K value for clustering
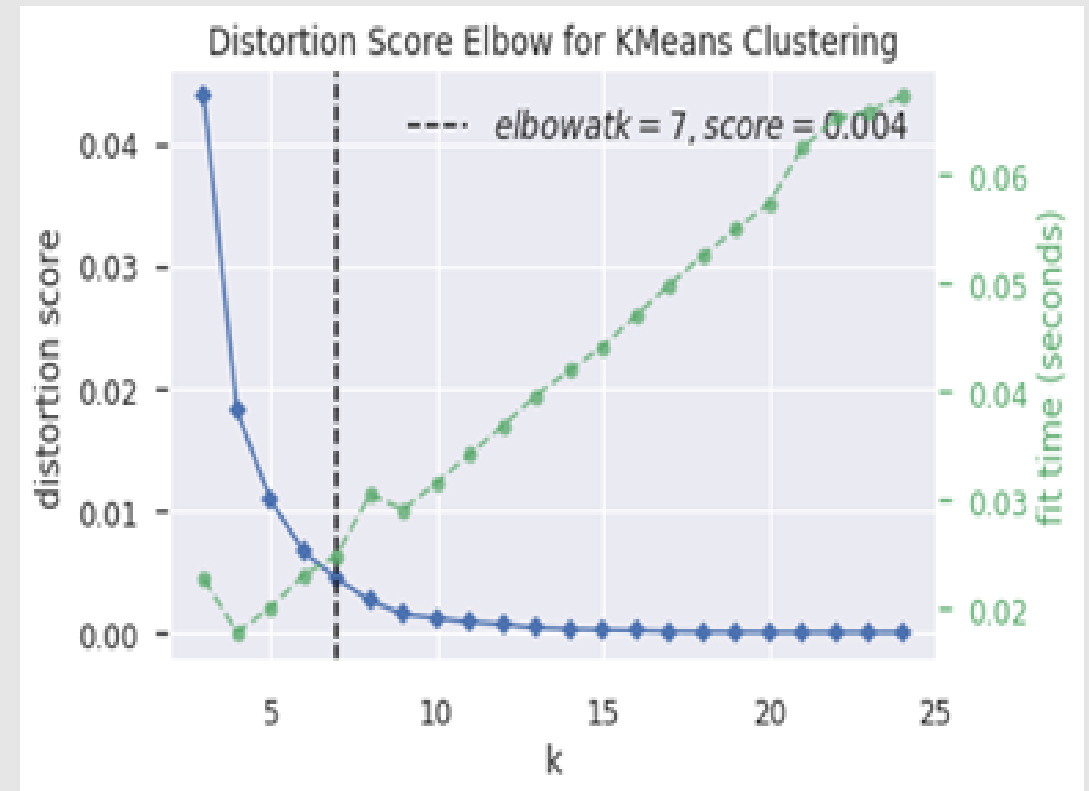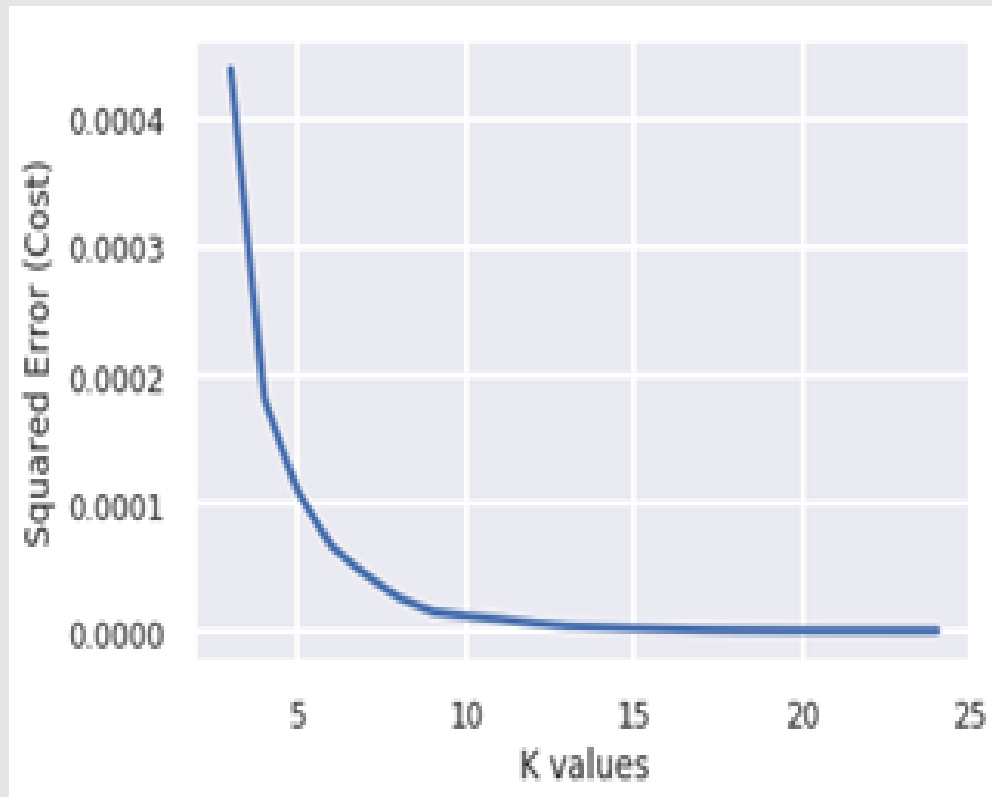
Methodology

◦ **Using**
  ◦ **Cluster Maps**
  ◦ **Individual cluster analysis**
  to find least populated neighbourhoods with pizza places and bars

◦ **Identifying locations for chain of gourmet pizza place with bar**

Results and Discussion

- K-means clusters map for pizza places
  - The red dots represent cluster 0 and purple dots represent cluster 1 with least number of pizza places
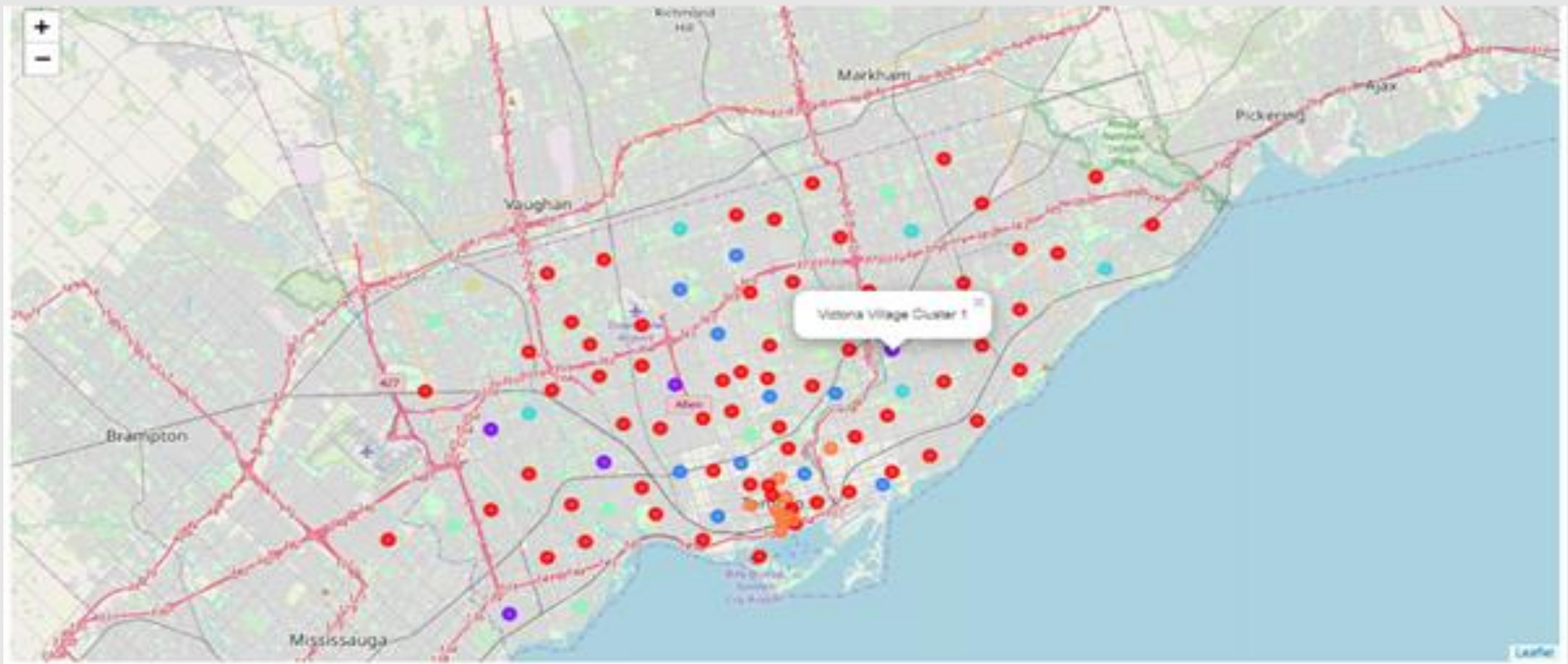
Results and Discussion

○ **K-means clusters map for bars**

  ○ **The red dots represent cluster 0 and purple dots represent cluster 1 with least number of bars**

Results and Discussion

○ **Cluster Zero: Clusters Analysis to find out neighbourhoods with least number of pizza places (i.e. North York, Scarborough)**

	North York	M3J	Flemingdon Park, Don Mills South	43.725869	-79.340923	0.0	0.0
62	North York	M3J	Northwood Park, York University	43.767980	-79.487262	0.0	0.0
63	North York	M3K	CFB Toronto, Downsview East	43.737473	-79.464763	0.0	0.0
64	North York	M3L	Downsview West	43.739015	-79.506944	0.0	0.0
65	North York	M3M	Downsview Central	43.728496	-79.495697	0.0	0.0
66	North York	M3N	Downsview Northwest	43.761631	-79.520999	0.0	0.0
69	North York	M6A	Lawrence Heights, Lawrence Manor	43.718518	-79.464763	0.0	0.0
71	North York	M6L	Downsview, North Park, Upwood Park	43.713756	-79.490074	0.0	0.0
73	North York	M9M	Emery, Humberlea	43.724766	-79.532242	0.0	0.0
74	Queen's Park	M9A	Queen's Park	43.667856	-79.532242	0.0	0.0
75	Scarborough	M1B	Rouge, Malvern	43.806686	-79.194353	0.0	0.0
76	Scarborough	M1C	Highland Creek, Rouge Hill, Port Union	43.784535	-79.160497	0.0	0.0
78	Scarborough	M1G	Woburn	43.770992	-79.216917	0.0	0.0
79	Scarborough	M1H	Cedarbrae	43.773136	-79.239476	0.0	0.0
80	Scarborough	M1J	Scarborough Village	43.744734	-79.239476	0.0	0.0
81	Scarborough	M1K	East Birchmount Park, Ionview, Kennedy Park	43.727929	-79.262029	0.0	0.0
82	Scarborough	M1L	Clairlea, Golden Mile, Oakridge	43.711112	-79.284577	0.0	0.0
83	Scarborough	M1M	Cliffcrest, Cliffside, Scarborough Village West	43.716316	-79.239476	0.0	0.0
84	Scarborough	M1N	Birch Cliff, Cliffside West	43.692657	-79.264848	0.0	0.0
85	Scarborough	M1P	Dorset Park, Scarborough Town Centre, Wexford ...	43.757410	-79.273304	0.0	0.0
86	Scarborough	M1R	Maryvale, Wexford	43.750072	-79.295849	0.0	0.0
87	Scarborough	M1S	Agincourt	43.794200	-79.262029	0.0	0.0
89	Scarborough	M1V	Agincourt North, L'Amoreaux East, Milliken, St...	43.815252	-79.284577	0.0	0.0

Results and Discussion

◦ **Cluster Zero: Clusters Analysis to find out neighbourhoods with least number bars (i.e. Central Toronto, Downtown Toronto)**

	Borough	Postcode	Neighborhood	Latitude	Longitude	Cluster Labels	Pizza Place	Bar
0	Central Toronto	M4N	Lawrence Park	43.728020	-79.388790	0.0	0.0	0.000000
1	Central Toronto	M4P	Davisville North	43.712751	-79.390197	0.0	0.0	0.000000
2	Central Toronto	M4R	North Toronto West	43.715383	-79.405678	0.0	0.0	0.000000
4	Central Toronto	M4T	Moore Park, Summerhill East	43.689574	-79.383160	0.0	0.0	0.000000
6	Central Toronto	M5N	Roselawn	43.711695	-79.416936	0.0	0.0	0.000000
7	Central Toronto	M5P	Forest Hill North, Forest Hill West	43.696948	-79.411307	0.0	0.0	0.000000
9	Downtown Toronto	M4W	Rosedale	43.679563	-79.377529	0.0	0.0	0.000000
12	Downtown Toronto	M5A	Harbourfront	43.654260	-79.360636	0.0	0.0	0.000000
14	Downtown Toronto	M5C	St. James Town	43.651494	-79.375418	0.0	0.0	0.000000
15	Downtown Toronto	M5E	Berczy Park	43.644771	-79.373306	0.0	0.0	0.000000
16	Downtown Toronto	M5G	Central Bay Street	43.657952	-79.387383	0.0	0.0	0.023810
21	Downtown Toronto	M5S	Harbord, University of Toronto	43.662696	-79.400049	0.0	0.0	0.055556
23	Downtown Toronto	M5V	CN Tower, Bathurst Quay, Island airport, Harbo...	43.628947	-79.394420	0.0	0.0	0.062500
26	Downtown Toronto	M6G	Christie	43.669542	-79.422564	0.0	0.0	0.000000
27	Downtown Toronto	M7A	Queen's Park	43.662301	-79.389494	0.0	0.0	0.025641
28	East Toronto	M4E	The Beaches	43.676357	-79.293031	0.0	0.0	0.000000
30	East Toronto	M4L	The Beaches West, India Bazaar	43.668999	-79.315572	0.0	0.0	0.000000
31	East Toronto	M4M	Studio District	43.659526	-79.340923	0.0	0.0	0.024390
34	East York	M4C	Woodbine Heights	43.695344	-79.318389	0.0	0.0	0.000000
35	East York	M4G	Leaside	43.709060	-79.363452	0.0	0.0	0.000000

Results and Discussion

- **The neighbourhoods with least number bars and pizza places and ideal for opening chain of gourmet pizza place with bar**

    - **Scarborough**
    - **North York**
    - **Central Toronto**

- **Future Work**
    - **Analysis of other factors that can affect this decision such as population, income per capita etc.**

Conclusion

- Firstly, the business problem was identified.

- Then, required data to solve the business problem was specified.

- The data was extracted from the identified sources and prepared for analysis.

- Exploratory data analysis was performed to have rough idea of tackling the business problem.

- Lastly, Machine learning i.e. k-means clustering was used to make decision to solve the business problem i.e. the neighbourhoods ideal for opening chain of gourmet pizza place with bar