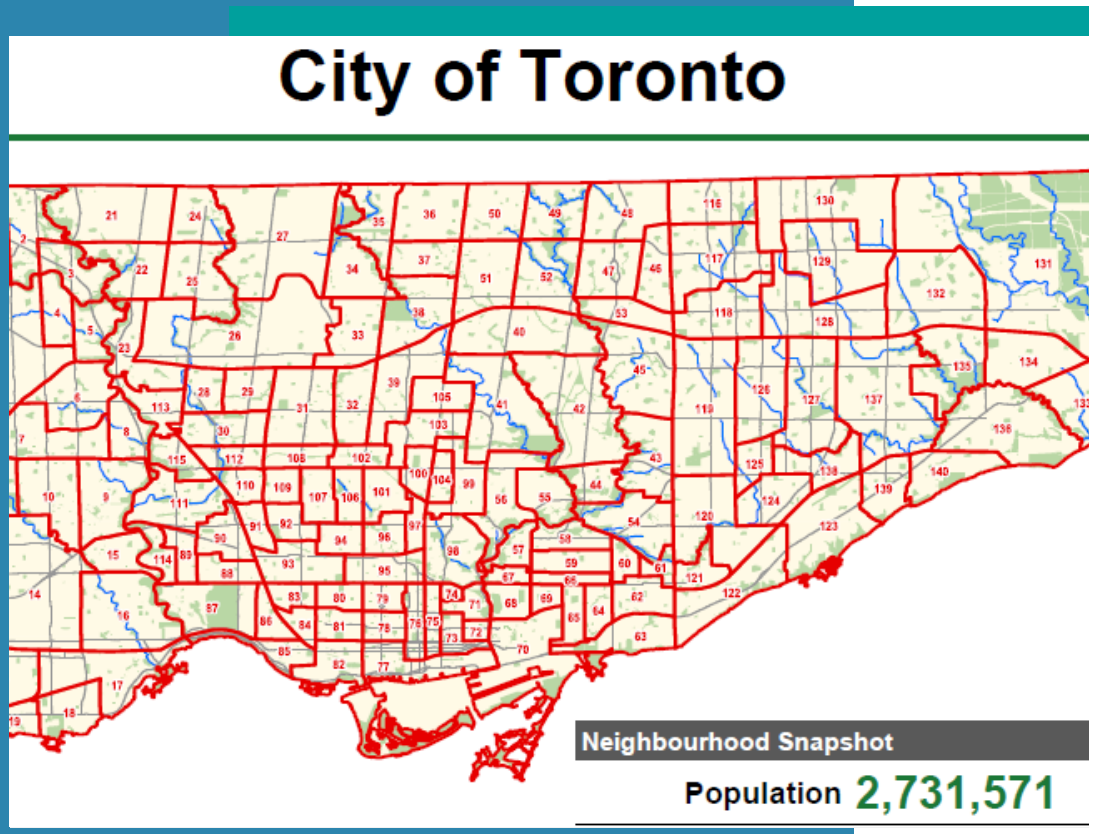
Opening an Italian Restaurant in Toronto

IBM Capstone Project presentation

By Chara Psoraki



Introduction

- Toronto is an **international center for business and finance.** It is also a **cultural center** with major Universities, and many arts and sports activities. This vibrant city provides many **opportunities** for young and enthusiastic individuals to open their own business and become a piece of this multidimensional city.
- Toronto's **diversity** is reflected in Toronto's ethnic neighborhoods, which include Chinatown, Corso Italia, Greektown, Kensington Market, Koreatown, Little India, Little Italy, Little Jamaica, Little Portugal and Roncesvalles (Polish community).

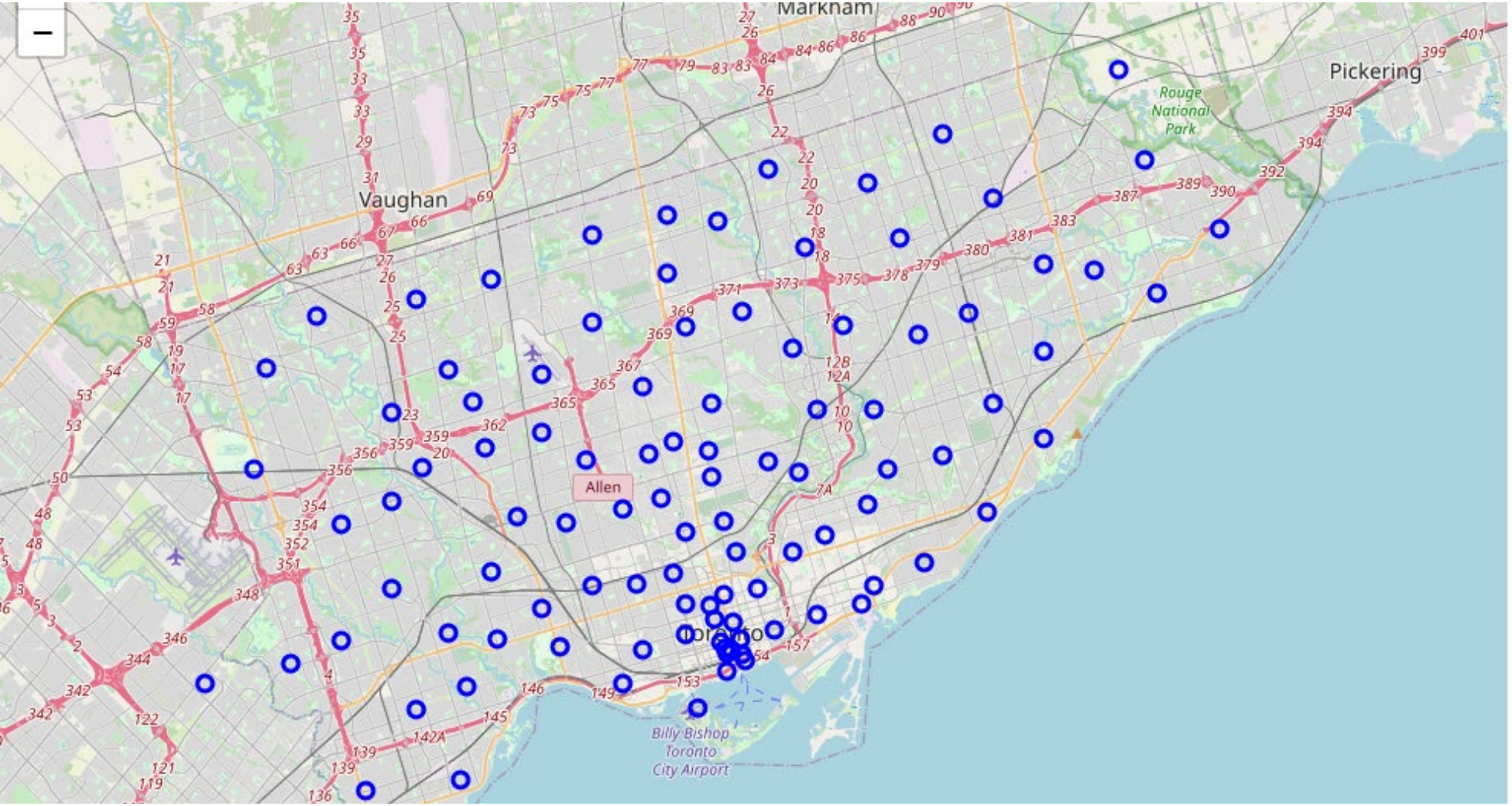
Opening an Italian Restaurant in Toronto
What would be the ideal location ?

Data collection

- The demographic data of the general population of Toronto were collected by the 2016 Canada Census
- The postal codes and location information for the neighborhoods of Toronto we collected from Wikipedia
- Data such as: Population, Average Family Income, Age, Gender for each neighborhood of Toronto were collected from Wellbeing Toronto's records
- Location information and venue information were extracted using the **Foursquare API**

Data analysis

Mapping the 103 Neighborhoods in Toronto



Using the Foursquare API and the GetNearBy function, we collected the location of 2,153 venues that belong to **277 unique categories**.

Group rows by neighborhood and by taking the mean of the frequency of occurrence of each category

```
toronto_grouped = toronto_onehot.groupby('Neighborhood').mean().reset_index()
toronto_grouped
```

	Neighborhood	Yoga Studio	Accessories Store	Adult Boutique	Airport	Airport Food Court	Airport Gate	Airport Lounge	Airport Service	Airport Terminal	American Restaurant	Antique Shop	Aquarium
0	Agincourt	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00
1	Alderwood, Long Branch	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00
2	Bathurst Manor, Wilson Heights, Downsview North	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00
3	Bayview Village	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00
4	Bedford Park, Lawrence Manor East	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.040000	0.000000	0.00
5	Berczy Park	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00
6	Birch Cliff,	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00

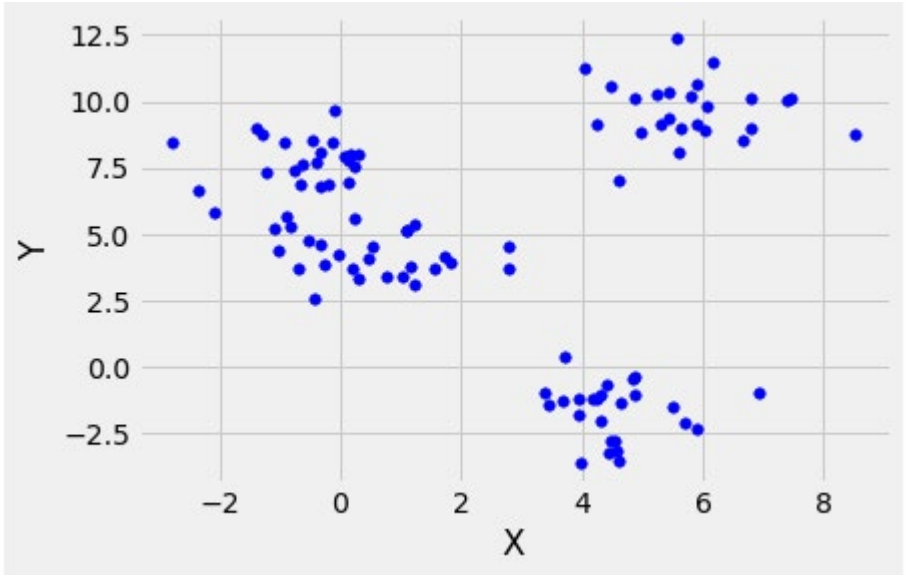
Clustering

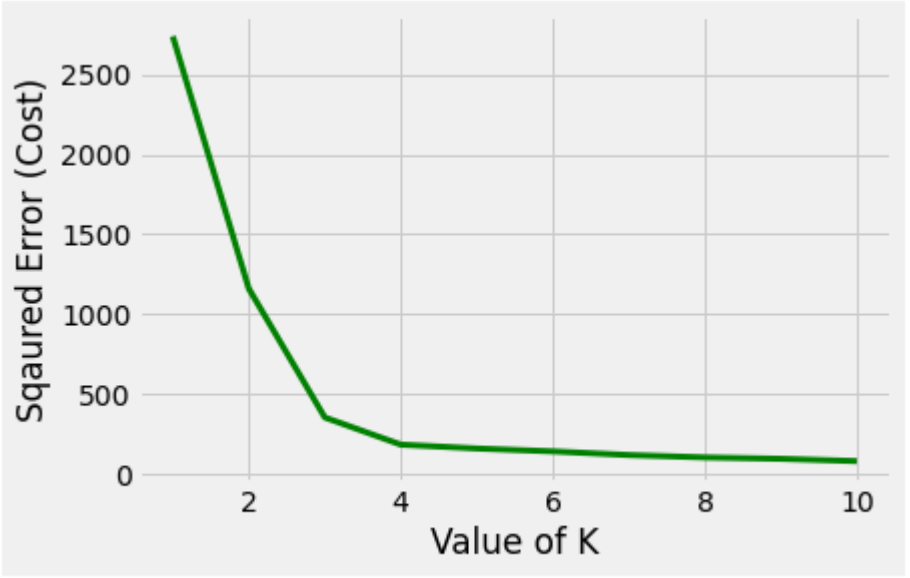
We created a new dataframe sorted by the “Italian Restaurants” . We cluster the neighborhoods with an initial K-value of 5.

	Neighborhood	Italian Restaurant	Cluster Labels
0	Agincourt	0.000000	0
1	Alderwood, Long Branch	0.000000	0
2	Bathurst Manor, Wilson Heights, Downsview North	0.000000	0
3	Bayview Village	0.000000	0
4	Bedford Park, Lawrence Manor East	0.083333	4
5	Berczy Park	0.017544	2
6	Birch Cliff, Cliffside West	0.000000	0
7	Brockton, Parkdale Village, Exhibition Place	0.043478	3
8	CN Tower, King and Spadina, Railway Lands, Har...	0.000000	0
9	Caledonia-Fairbanks	0.000000	0
10	Cedarbrae	0.000000	0
11	Central Bay Street	0.044118	3
12	Christie	0.066667	4
13	Church and Wellesley	0.000000	0

Clustering

Finding the best K-value with samples generator and calculating the cost (squared error)

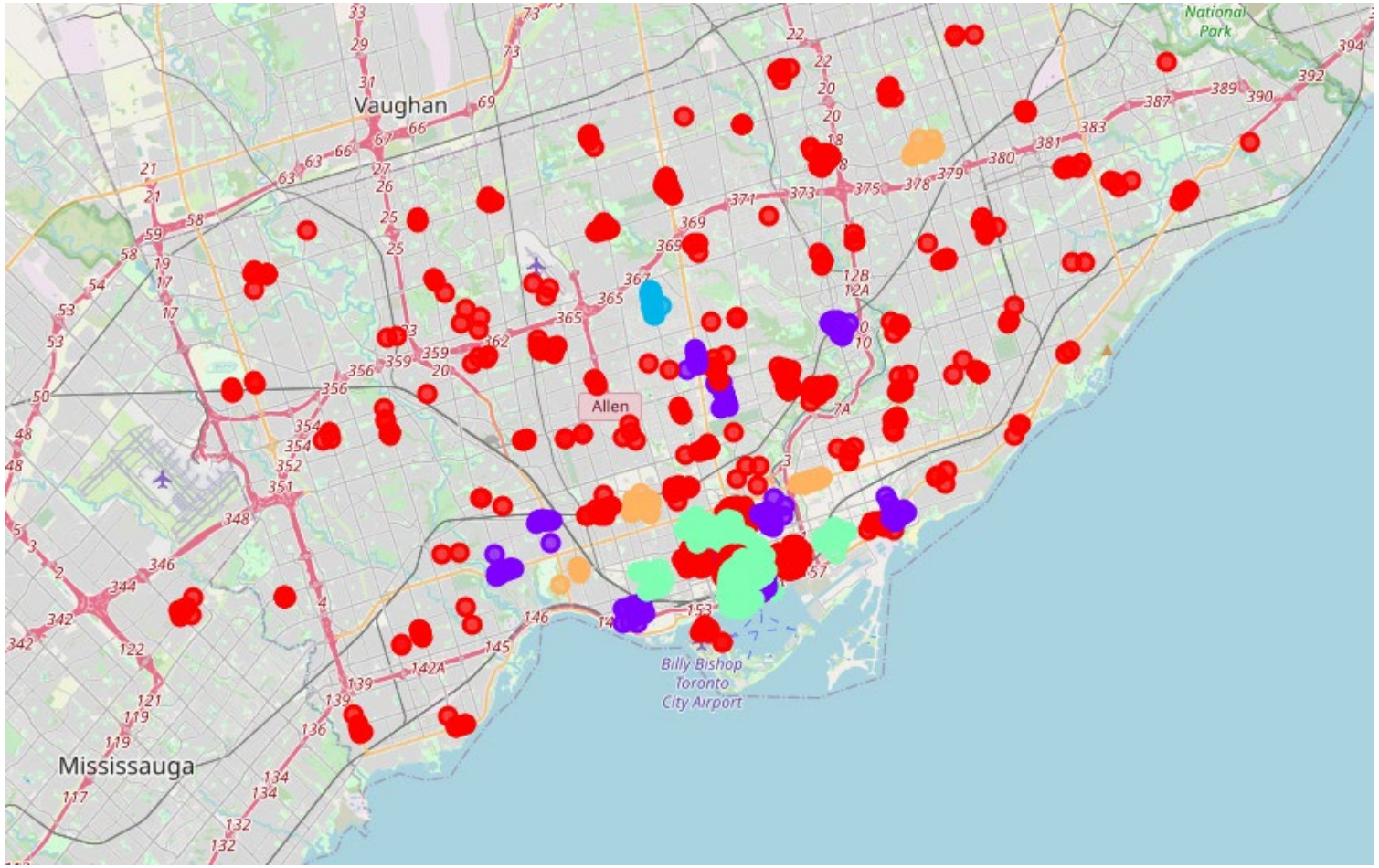




The **best K-value is equal to 4** because from that point and after we have the least error.

Clustering

Running again the cluster analysis for K=4



The map with the 4 clusters

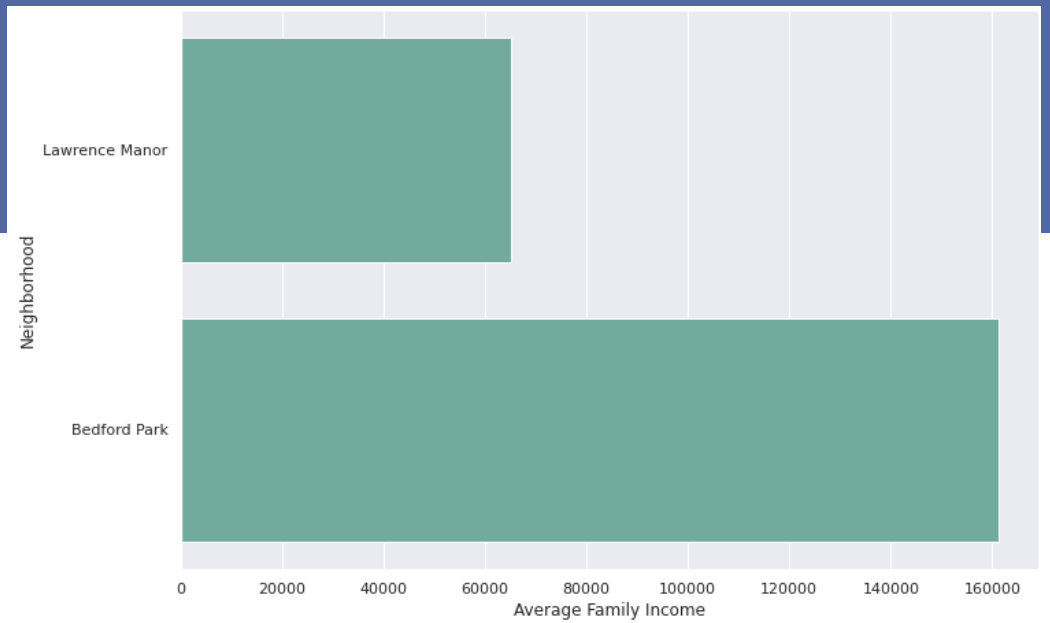
Clustering

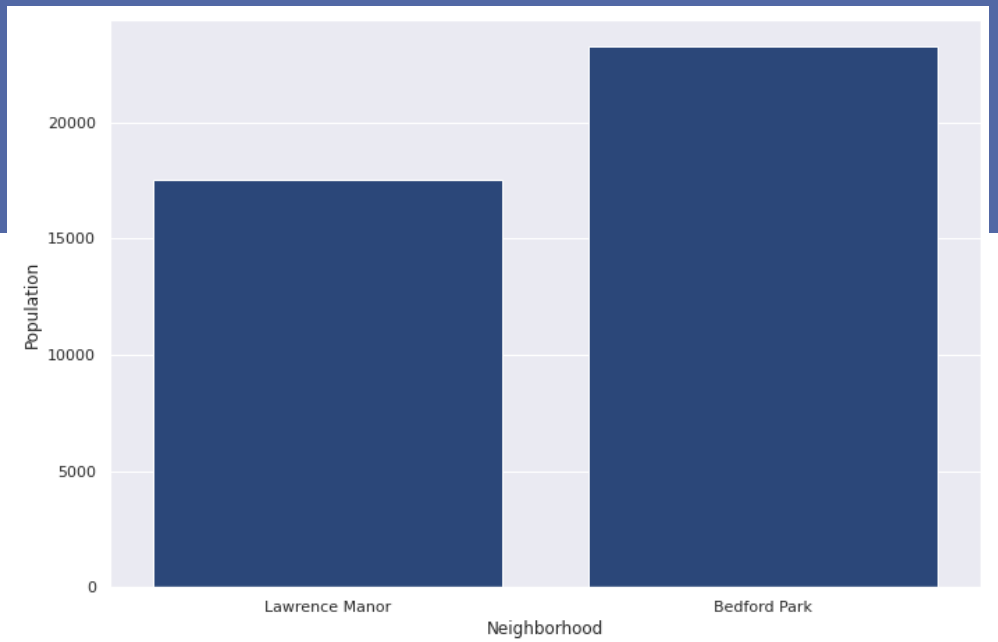
Total venues in neighborhoods in clusters:

Total venues in cluster 0 = 945
Total venues in cluster 1 = 396
Total venues in cluster 2 = 28
Total venues in cluster 3 = 689

The 3rd group (cluster 2) has the lowest concentration of venues (1,3%).

Analyzing the 3rd cluster





Average Family Income

- **Bedford Park** 's Average Family income is more than double compared to the one of Lawrence Manor.
- Bedford has significantly higher Average Family income ($ 161,100) compared to the Toronto's Median Family income ($65,829).

Total population

- The population of Bedford Park is also significantly higher (24.6%) that the one at Lawrence Manor.
- A higher population is a good indicator of higher demand in restaurants and shops

Conclusions

Based on the cluster analysis it is advise to open an Italian restaurant at the third cluster that is consisted by the neighborhoods of Lawrence Manor and Bedford. This cluster has **only 28 restaurants** in these neighborhoods, thus less competition. Both neighborhoods are located outside the big city center of Toronto.

Bedford has significantly higher Average Family income ($ 161,100) compared to the Toronto's Median Family income ($65,829). Therefore, Bedford is a neighborhood that you can open an Italian restaurant **with less risk.**

Future directions

- Rent prices and how they influence the choice of location
- How touristic are the Toronto neighborhoods. How does that influence the restaurant’s traffic
- Ethnicities and cultures of each neighborhood and their preference in international cuisines