

## The Battle of Neighborhoods – Toronto Restaurant

### Business Problem

The idea of this study is to leverage the foursquare location data to help client identify the optimal location to open a new restaurant in Toronto. To do this I am going to look into proving data about income and population size of each neighborhood and the competitors that already operated in the area.

### Data

In order to provide the client all necessary information, I'll be online and available data that contains all information I needed (population, income per household and competitors in the area by using Foursquare API ).

Here is publicly available data:

- <https://www.toronto.ca/city-government/data-research-maps/open-data/open-data-catalogue/#8c732154-5012-9afe-d0cd-ba3ffc813d5a>
- <https://www.toronto.ca/city-government/data-research-maps/open-data/open-data-catalogue/#a45bd45a-ed8-730e-1abc-93105b2c439f>
- [https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada:\\_M](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)

### Methodology

For this project I had used several data set which mentioned above. 1) Wikipedia and 2) Foursquare API to find out competitor. These two data contain valuable information such as income, population and competitors which help client decide the best neighborhood to open a new restaurant.

Additionally, I use K means clustering algorithm and explore the potential areas by creating several cluster of locations.

### Results

Our analysis shows that there are number of restaurants in Toronto but there are small number of resultant close to the center. We found 4 boroughs and 74 neighborhoods inside geographical coordinate of 43.653963, -79.387207. Our Cluster 4 and Cluster 5 shows higher number of restaurants.

### Conclusion

This report is helpful for client whom is planning to open a new restaurant in Toronto by comparing different neighbourhoods in Toronto. Foursquare data identified boroughs that justify further analysis. Clustering locations was performed to create major zones of interest. However, it may not cover all variables so it can't be used as a single decision-making tool.