

The background features a large, abstract graphic element in the upper right corner composed of several overlapping triangles in various shades of green, from light lime to dark forest green.

**Capstone Project - Where is the best  
place to open a restaurant in Toronto?**

# Introduction/Business Problem

- ▶ Before opening a new restaurant in any place, it is essential to investigate the demand for it.
- ▶ It is crucial to investigate the criteria such as population density, availability of other restaurants (competition), income diversity of people around that area to identify whether opening a restaurant would be profitable.
- ▶ These features collectively enable identifying the hotspots for a restaurant that would generate the best income. In this project, we consider these aspects in identifying the best location in Toronto to open a restaurant by a certain client.
- ▶ **The business problem is identifying and recommending the client with a certain location/s in Toronto that would produce the best profit.**
- ▶ In order to achieve this task, we will investigate the population density, availability of other restaurants, and income diversity of the people in the Toronto neighborhoods.

# Data

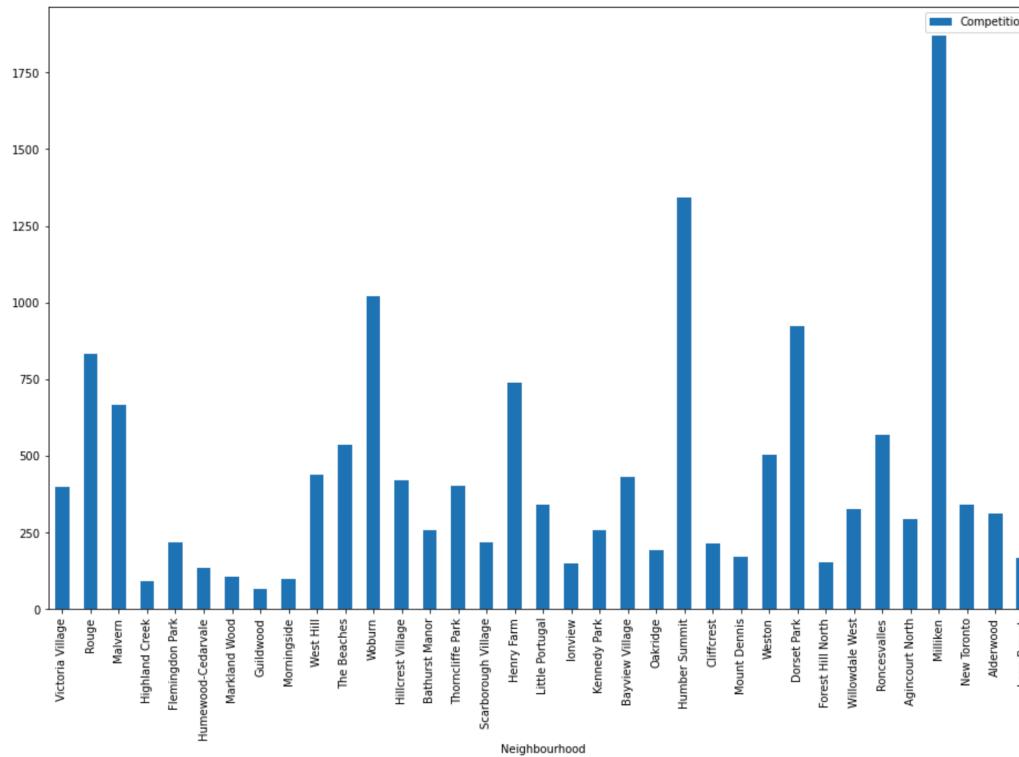
- ▶ The project used the data available in the following links
- ▶ [https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada:\\_M](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)
- ▶ [https://github.com/courseraassignmentcapstone/Coursera\\_Capstone/blob/main/Geospatial\\_Coordinates.csv](https://github.com/courseraassignmentcapstone/Coursera_Capstone/blob/main/Geospatial_Coordinates.csv)
- ▶ <https://foursquare.com/>
- ▶ <http://map.toronto.ca/wellbeing/#eyJ0b3Itd2IkZ2V0LWNsYXNzYnJlYWsiOsSACGVyY2VudE9wYWNpdHnElzcwfSwiY3VzxIjtYcSTYcSXxIBuZWlnaGJvdXJob29kc8S2fcSrxIHeg8SFxIfEicSLdGFixYXEmCLEo3RpdmVUxZBJZMSXxYnEhMWPYi1pbmRpY2HEgnLFhcWIYWdzTWFwxLYiesWCbcSXMTPErHjEly04ODM3NzYzLjXGhDcyN8SsxKc6NTQzMjkzMS4yNMaDMjg1xYjFpMWmxajFqsWSxIDFmMWraW9uxJcyxKxzxarnbGXFhsSsxZZtZXPEm2nGtsayxK3Ev8STxJ9JxaXFp8WpxINNxYPGsToixq1uxq%2FGscWH>

# Methodology

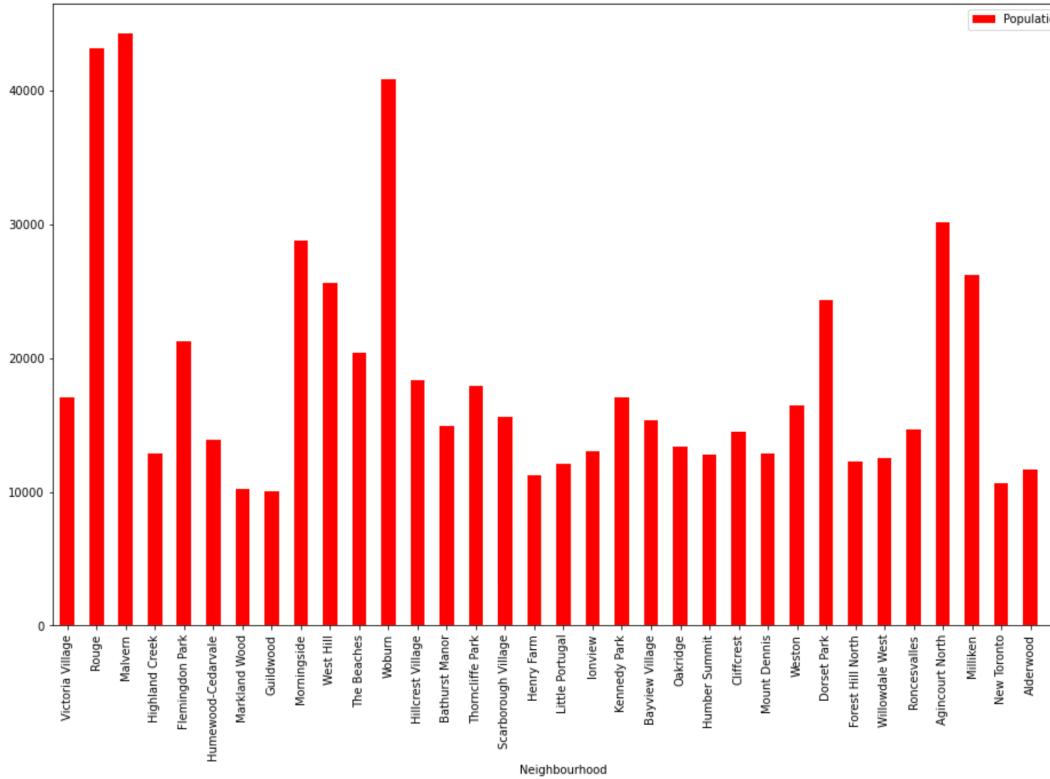
- ▶ We first conducted a basic descriptive analysis of the datasets collected to get a brief idea of the dataset (collected from [https://github.com/courseraassignmentcapstone/Coursera\\_Capstone/blob/main/toronto\\_data.csv](https://github.com/courseraassignmentcapstone/Coursera_Capstone/blob/main/toronto_data.csv)) characteristics related to the project.
- ▶ This project aims to conduct a suitability analysis for a new restaurant to open in Toronto neighborhoods.
- ▶ We will limit our analysis to the top 100 venues that are in Toronto within a radius of ~8km. Next, we will run cluster analysis to identify the best place to open a restaurant.
- ▶ This analysis will plot the points and clusters in maps, generate bar plots, and conduct basic descriptive analysis to suggest the best neighborhoods to open a new restaurant.
- ▶ The stakeholders can then utilize this information to open a restaurant that produces a good income.

# Analysis, Results and Discussion

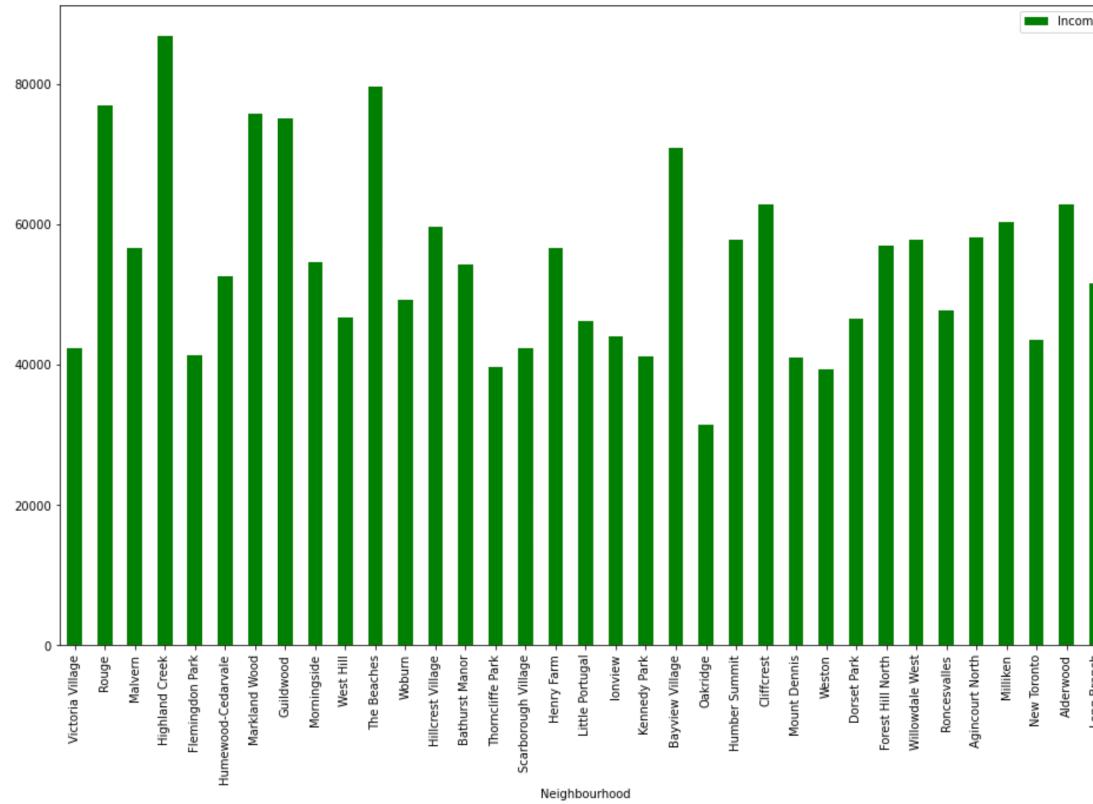
- We were able to obtain the following bar plots for the Competition, Population, and Income



# Analysis, Results and Discussion (cont'd.)



# Analysis, Results and Discussion (cont'd.)

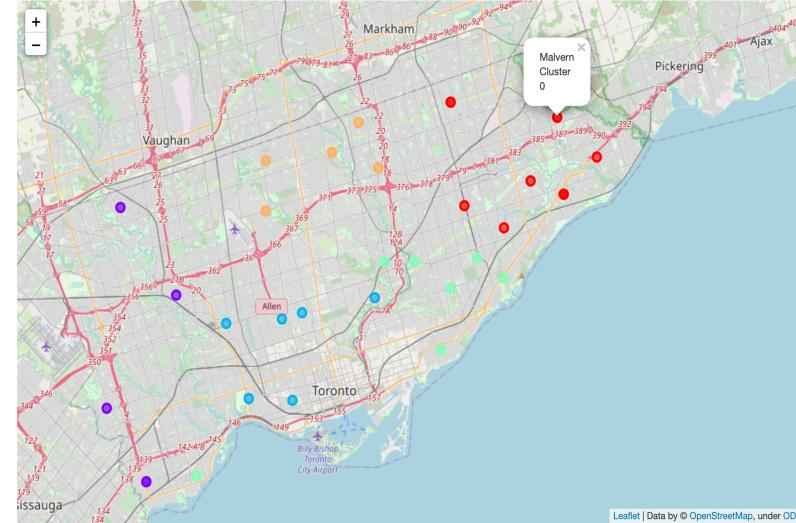
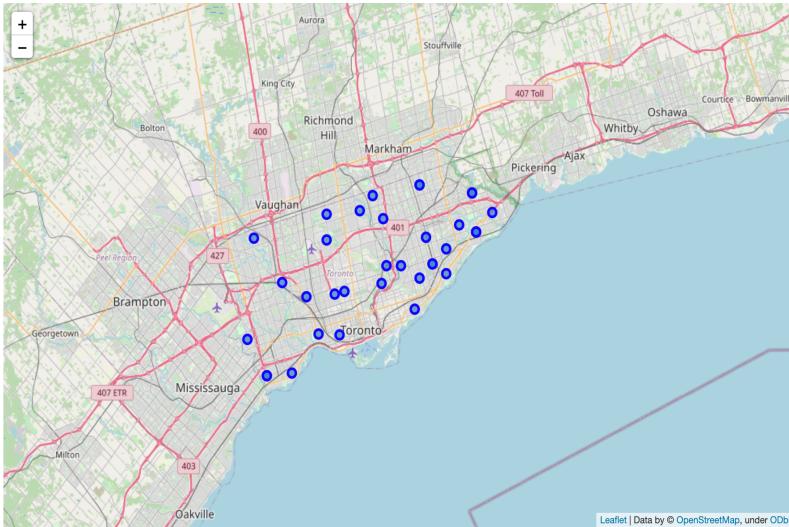


# Analysis, Results and Discussion (cont'd.)

- ▶ Milliken shows the highest competition for a new business as shown in the bar chart and the above record. This suggests that it is challenging to open a new business in this area.
- ▶ Malvern has the highest population. This means that a restaurant in Malvern may attract many customers.
- ▶ As shown in the bar graph and the above record, Highland Creek is the area with the highest income. This may have an impact on people to buy from restaurants regularly.

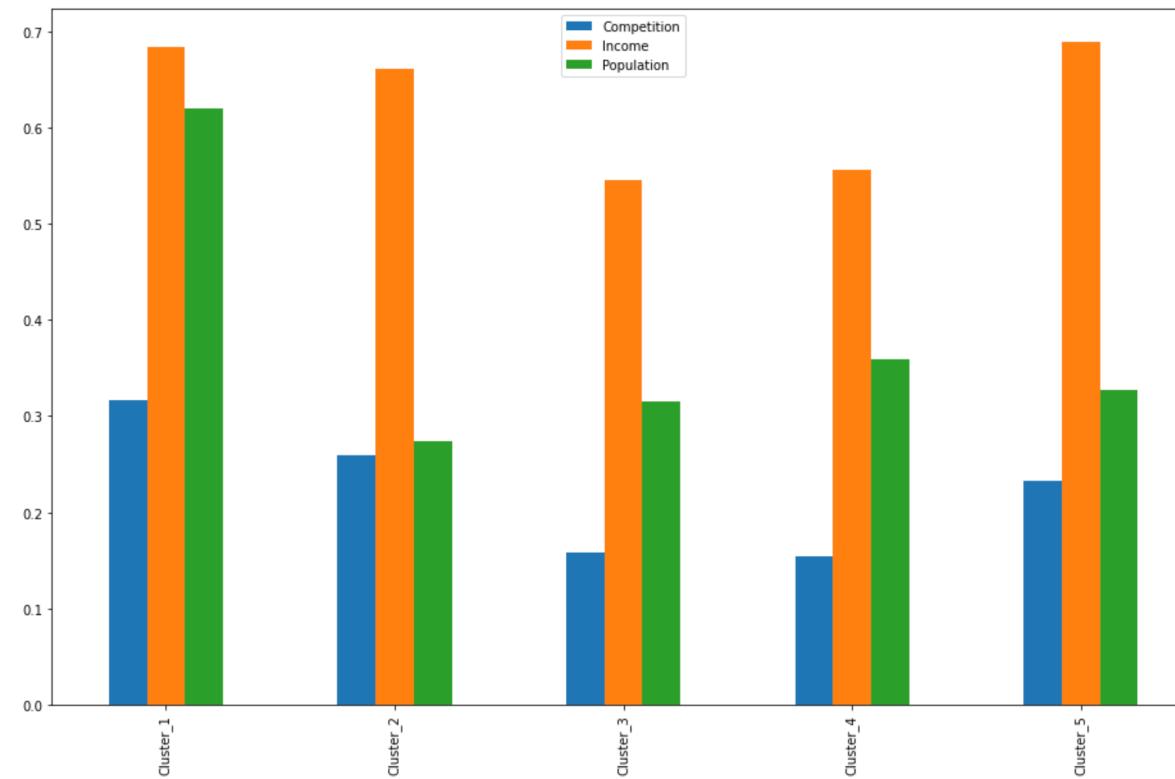
# Analysis, Results and Discussion (cont'd.)

- We can plot the original data and the cluster data on a map as follows



# Analysis, Results and Discussion (cont'd.)

- ▶ Next, we can investigate how the income varies in each cluster using line plots for each cluster



# Analysis, Results and Discussion (cont'd.)

- ▶ According to our analysis, the first cluster (numbered as cluster 0 on the map) provides the best place to start a new restaurant.
- ▶ According to the bar chart, it can be seen that there is a high population and income although the competition is a little bit high.
- ▶ However, the competition compared to other clusters is not drastically high in the first cluster.
- ▶ As it is visible, the competition is close to that of clusters 2 and 5.
- ▶ Hence, we suggest that the area around the Malvern (Malvern cluster 0 as depicted in the map) is the best place to open a restaurant.

# Conclusion

- ▶ In this project, we tried to identify a suitable place for a new restaurant based on three factors; competition, population, and income.
- ▶ We first extracted data on these statistics, and then we merged the geographical data to the corresponding data to conduct geographical analysis based on clustering.
- ▶ We showed the dynamics of different places in Toronto. Then we also suggested a certain area that shows the best performance in terms of competition, population, and income.
- ▶ However, the final decision to open a restaurant in a certain place will be taken by the stakeholders. The stakeholders can effectively use the insights produced in the project to take this decision effectively.