CAPSTONE PROJECT - THE BATTLE OF THE NEIGHBORHOODS (WEEK 2)

APPLIED DATA SCIENCE CAPSTONE BY IBM/COURSERA

BUSINESS PROBLEM

For an investor who would like to invest their money in the restaurant business, it may be important to know the level of competition in the area where he plans to open a restaurant.

Using open data, it is possible to determine the number of restaurants per 1000 people. I will be using Toronto's public statistics.

We will determine the population of the city districts. Then I will pull data from Foursquare to determine the number of restaurants. Using this data, you can determine the level of competition in each borough of Toronto.

This will allow the investor to more easily decide in which district of the city to open a restaurant

DATA ACQUISITION AND CLEANING

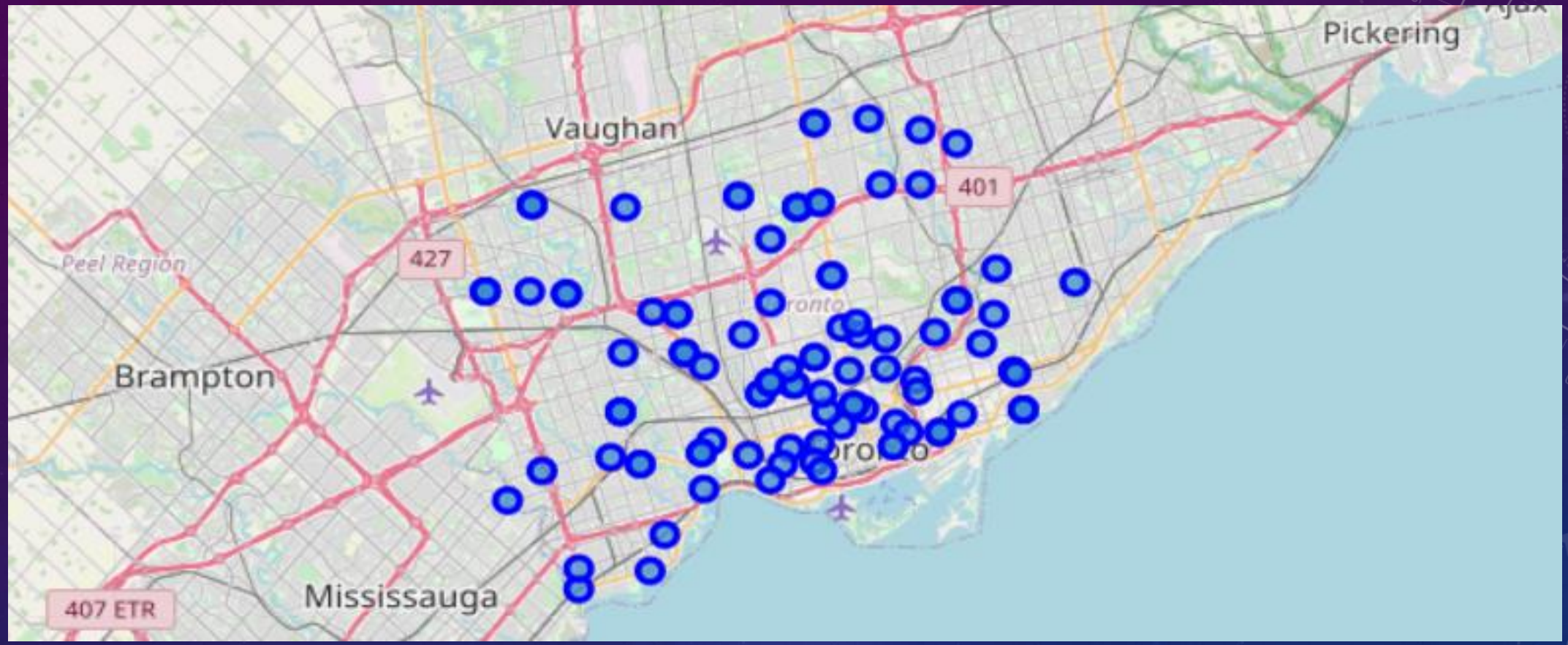
- Download the data from http://map.toronto.ca/wellbeing/ The downloaded file contains data about
- the neighbourhood
- the population

in these neighbourhood.

The Geopsy service was used to determine the coordinates of the city's districts.

Geopy is a Python client for several popular geocoding web services.

Geopy makes it easy for Python developers to locate the coordinates of addresses, cities, countries, and landmarks across the globe using third-party geocoders and other data sources.

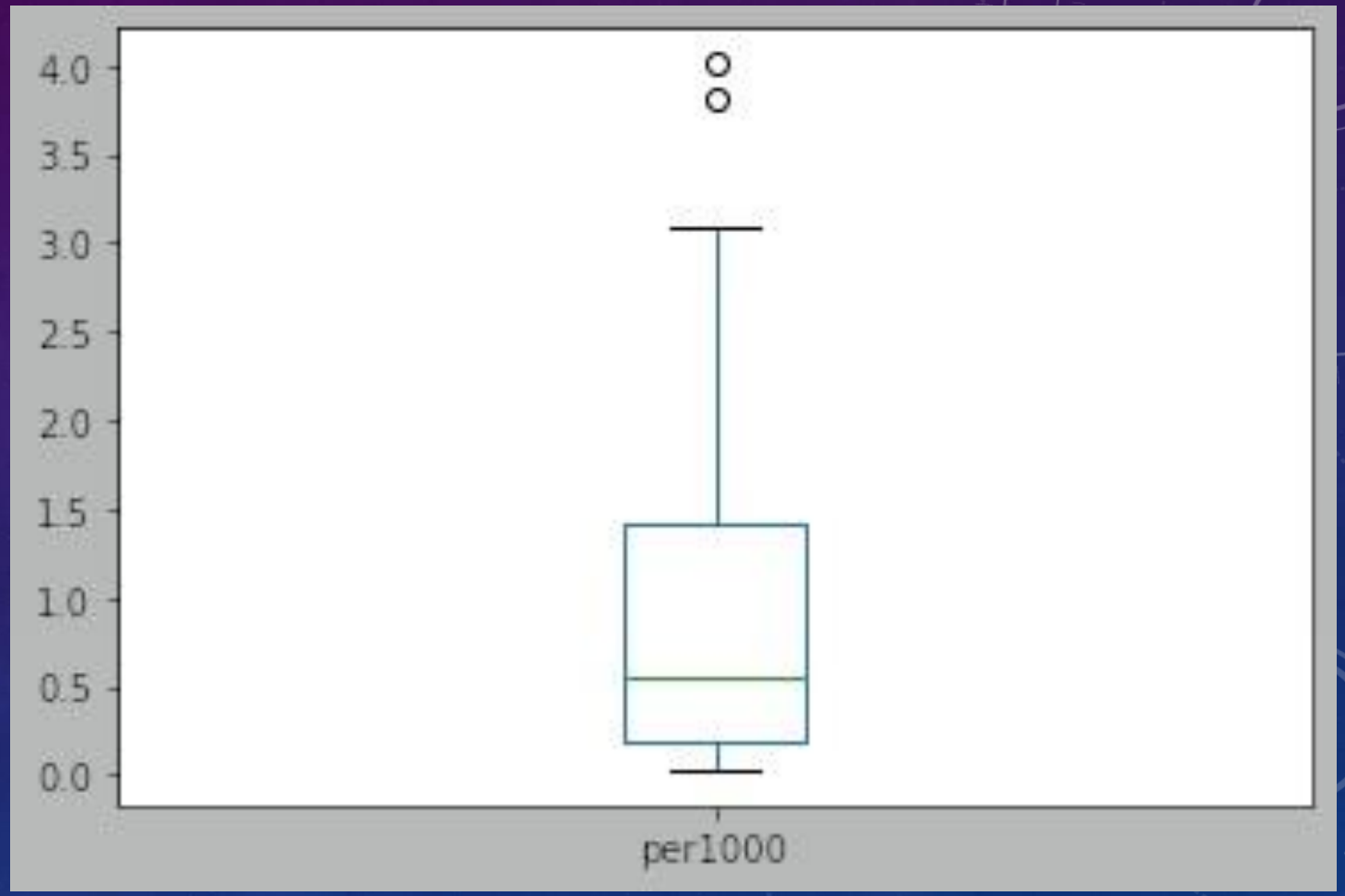


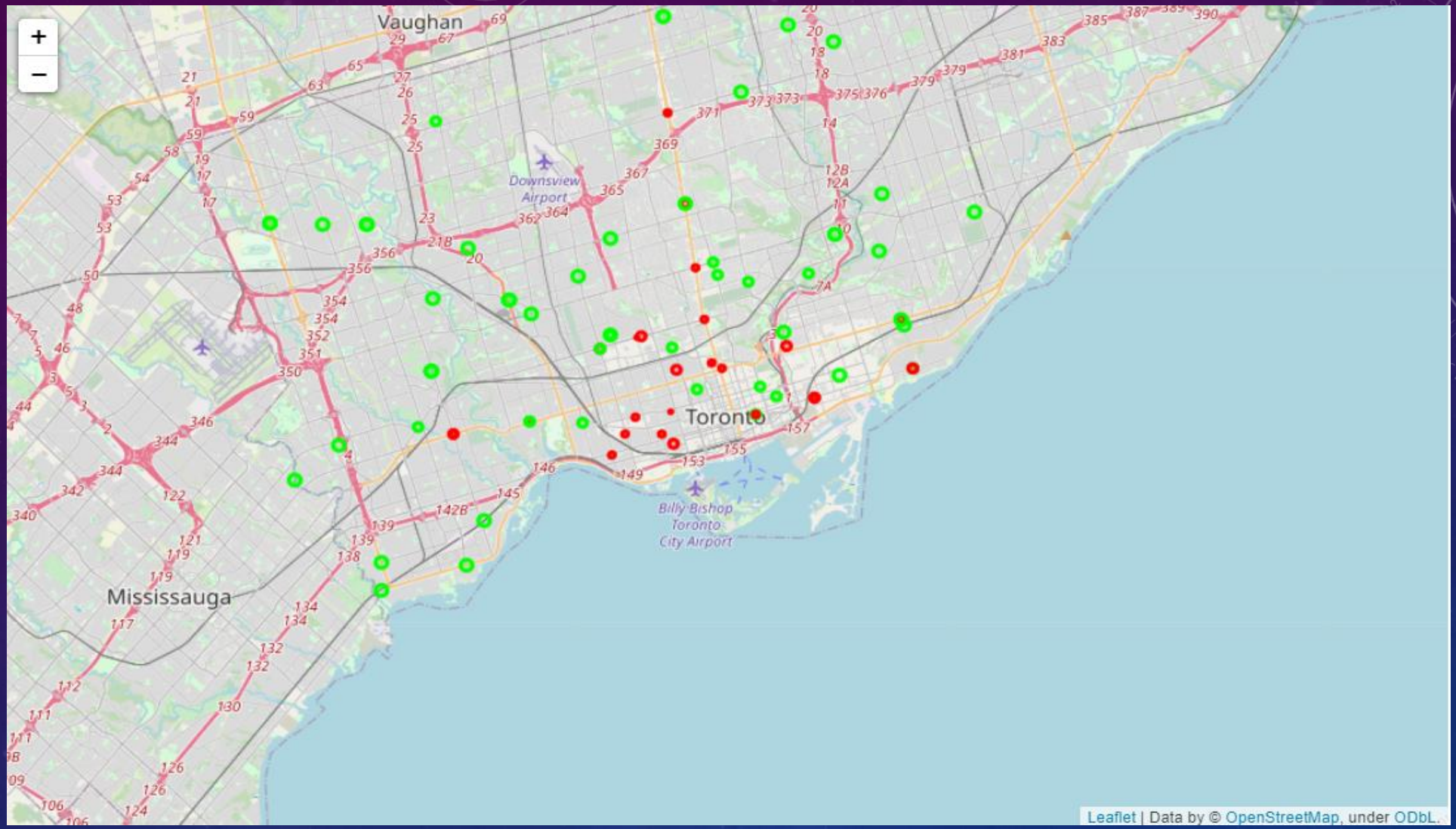
Neighbourhoods of Toronto

THE TOP 10 NEIGHBORHOOD WITH THE SMALLEST DENSITY OF RESTAURANTS PER 1000 PEOPLE.

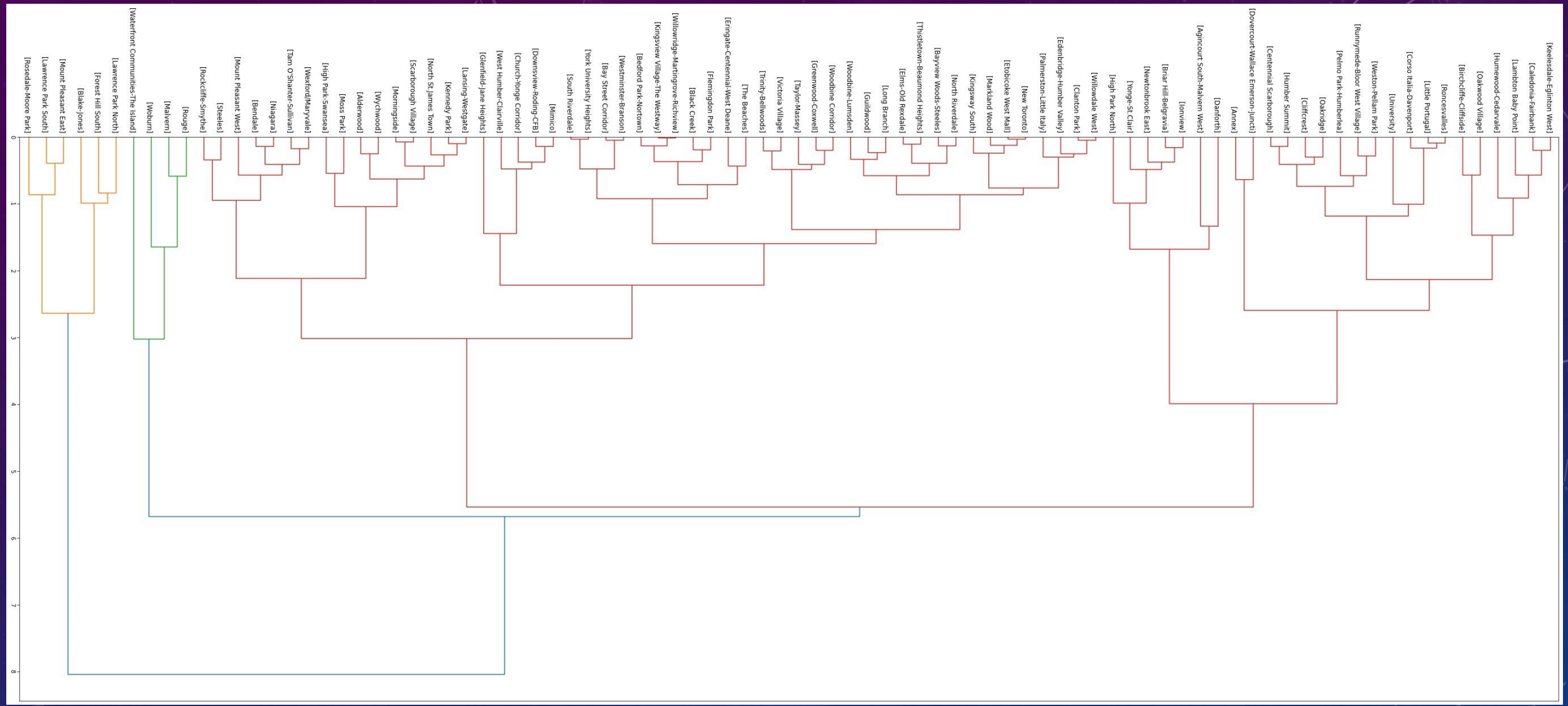
Neighborhood	Total Population	per1000
Mimico	33964.0	0.029443
Church-Yonge Corridor	31340.0	0.031908
Flemingdon Park	21933.0	0.045593
Downsview-Roding-CFB	35052.0	0.057058
Westminster-Branson	26274.0	0.076121
Bay Street Corridor	25797.0	0.077528
Etobicoke West Mall	11848.0	0.084402
New Toronto	11463.0	0.087237
Black Creek	21737.0	0.092009
Markland Wood	10554.0	0.094751

BOX PLOT OF THE NUMBER OF RESTAURANTS PER 1000 PEOPLE.





A map of Toronto neighbourhoods showing the number of restaurants per 1000 people.



Determine which areas of Toronto are most similar , that is, how you can group the areas, which areas are most similar to each other in terms of the number of restaurants per 100 people. Our goal here is to use clustering techniques to find the most distinctive groups of Toronto neighborhoods. This will help investors make a decision about opening a restaurant in a certain area.

CONCLUSION AND FUTURE DIRECTIONS

The aim of this project was to identify areas with the least number of restaurants per 1000 people to help investors in the restaurant business narrow down the search for the optimal place to open a new restaurant.

- By calculating the distribution of restaurant density based on Foursquare data, we have identified the 10 most attractive areas for investors.
- Clustering and visualization of districts showed areas with a high number of restaurants for every 1000 residents, in which it is not advisable to open new restaurants.
- We performed hierarchical clustering to determine which areas of Toronto are most similar , that is, how to group areas, which areas are most similar to each other in terms of the number of restaurants per 100 people.