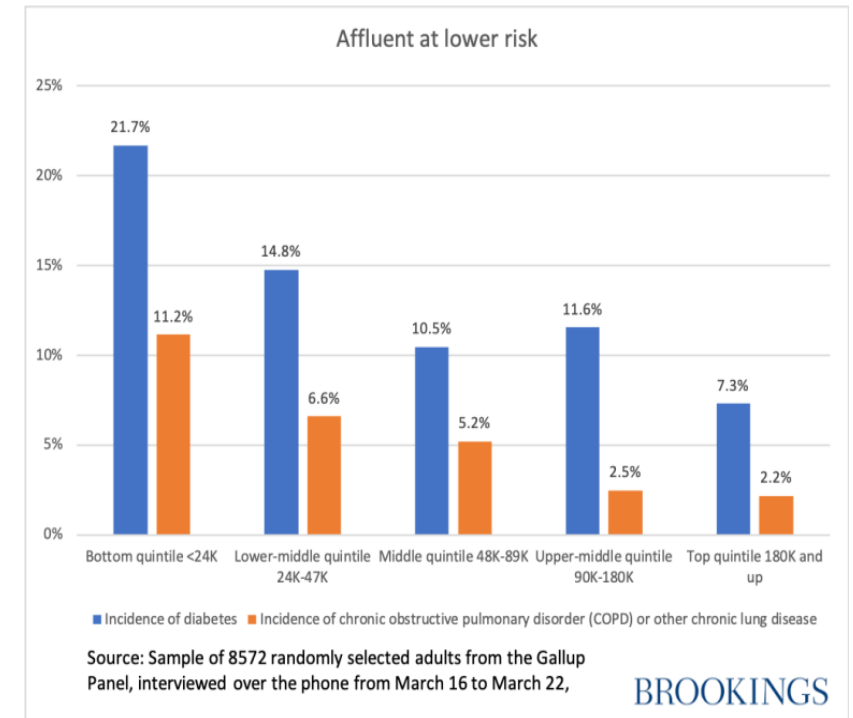# Analysis of Gyms and Sporting venues' distribution across Toronto

Salah Gaffar

Data Scientist

Why having enough exercise is especially a necessity now?

- Increasing number of deaths due to COVID-19, especially in countries where obesity levels and pre-existing conditions are alarming

- Canada ranks 44th out of 190 countries (26th when only males are considered) in terms of overall BMI scores with an average of 27.2 (above 25 is considered overweight)

- Due to the currently record healthcare costs, any relief in the existing burden on healthcare system will save lives

Affluent at lower risk

Source: Sample of 8572 randomly selected adults from the Gallup Panel, interviewed over the phone from March 16 to March 22,

BROOKINGS

Data acquisition and preparation

Data was obtained mainly from these sources:

- Foursquare API venue data

- Demographic data on each Toronto neighborhood (from 2006 census)

- **Toronto Wellbeing** web tool

Data from the variety of sources were cleaned, filtered and merged in order to obtain the derived variables to feed them into the machine learning model, which included:

- Average income for each neighborhood

- Number of venues per capita

Exploratory data analysis

- A wide income disparity is apparent in the average income across different neighborhoods in Toronto

- The general case for neighborhoods in Toronto is that there aren't enough sporting venues except for some outliers where enough venues exist.

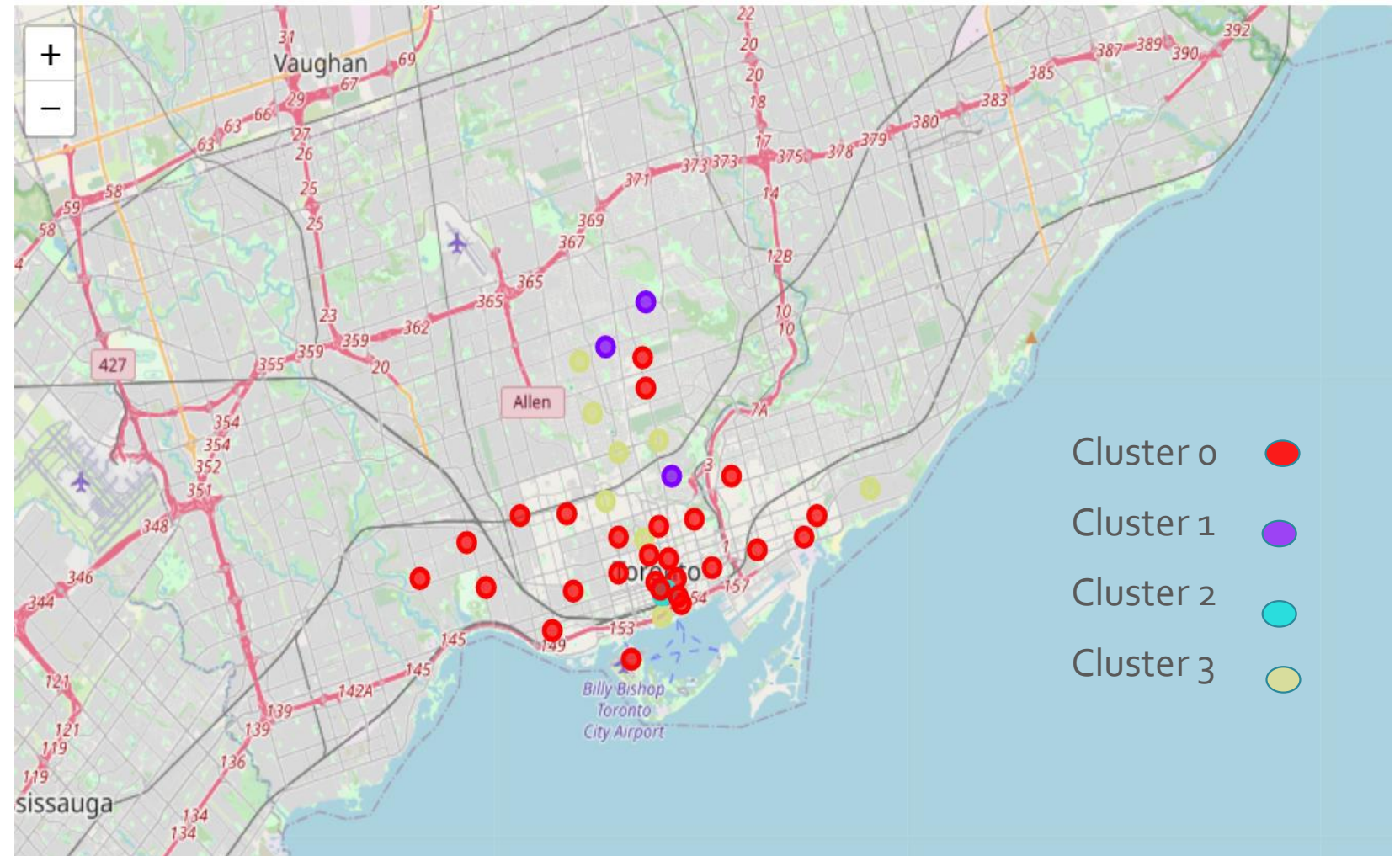Average income distribution across Toronto
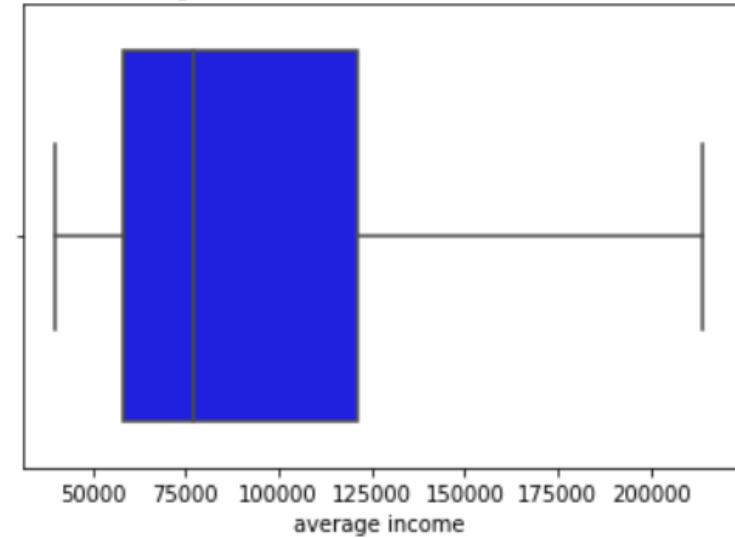

Venues per capita across Toronto
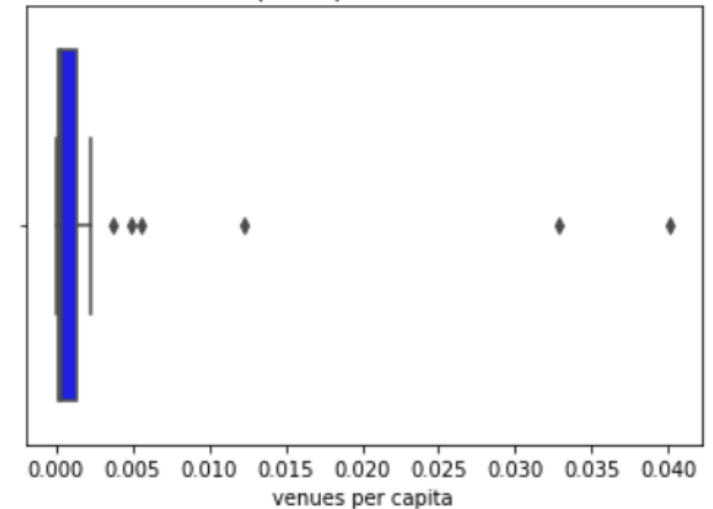
K-cluster distribution

The K-clusters were distributed as shown below:

K-Cluster Insights

- Mismatch between average income and availability of sporting venues means that business owners should target different neighborhoods according to its shown segment.

- Large sporting corporations should target neighborhoods with high income and low venue representation.

- Charities, NGOs and government should encourage small business owners to open small-scale sporting venues in clusters where income is low to medium

- Areas where enough venues exist should not intentionally targeted in terms of incentives for opening more sporting venues.

Comparison of average income among the 4 clusters

Comparison of venues per capita among the 4 clusters
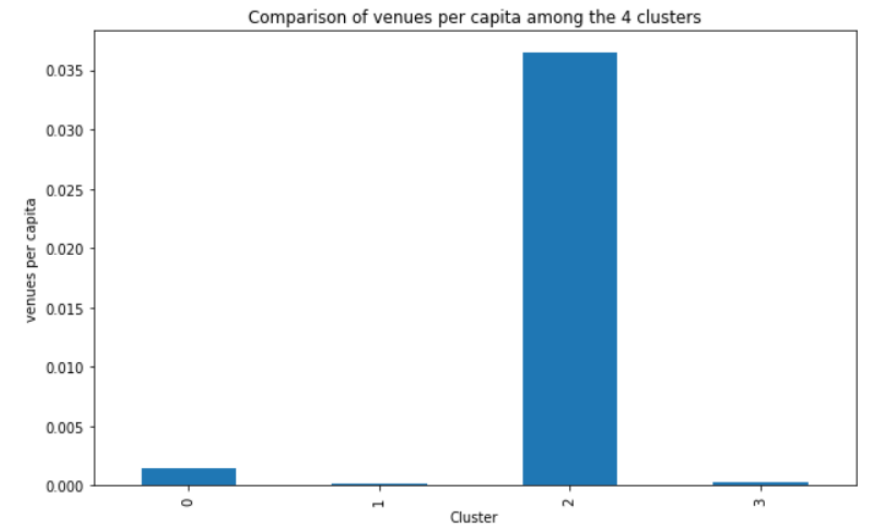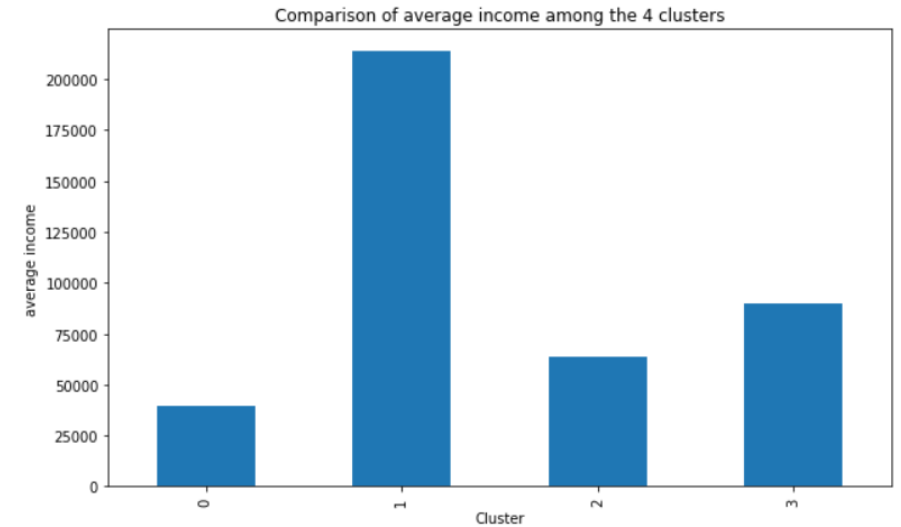
Conclusions and future work

The following additional considerations can be included:

- Classifying venues according to the recommended capacities for each venue (e.g. a skating rink can occupy more than a Yoga studio, so each venue should be weighted by its capacity to serve customers

- Parks and social clubs and other places with available sporting facilities can be included in this study

- Adding the real-estate prices and renting costs paid by businesses in each area and adding this data to the clustering inputs, so as to divide the prospect neighborhoods into different classes in terms of the costs incurred to the business owner who invests there as in addition to the availability of customers, which is already displayed in the current study.

- Use more recent census data, where the current dataset being used was last updated in 2006.

- Use more weighting parameters for venues to include how far they are from the neighborhood center, since our current 500m radius can be considered too much for elderly citizens that would like to find a place to exercise