NEW BUSINESS LOCATION IN TORONTO, CANADA

Problem statement

- ▶ The best location in Toronto for opening a club is to be recommended.
- ▶ Very specific not only boroughs but also neighborhoods.
- ▶ Lower competition, greater accessibility.

Initial Research Results

- ▶ The highest concentration of young people in East York.
- ► Scarborough in 20-minute drive from East York.
- ▶ Low competition, close to the busy parts.
- ▶ Which neighbors of Scarborough are popular with the youth?

Data and methodology

- We need all neighborhoods in Toronto with their longitudes and latitudes.
- ▶ We have data on the coordinates of Postal Codes, but as we will need to make the research on the neighborhoods and boroughs, we are to create the necessary database. Thus we go through the following steps:
 - ► First the list of the boroughs, neighborhoods and their Postal Codes will be extracted from the following page:

 https://en.wikipedia.org/wiki/List of postal codes of Canada: M
 - Observations with not assigned boroughs are dropped, repeating postal codes grouped, not assigned neighborhoods carry their borough name.
 - ▶ The data will be collated with the geospatial coordinates of the Postal Codes.
 - ▶ Referral feature (Postal Code) renamed to match the latter in the first dataset.
 - We finally merge these two datasets on Postal Code

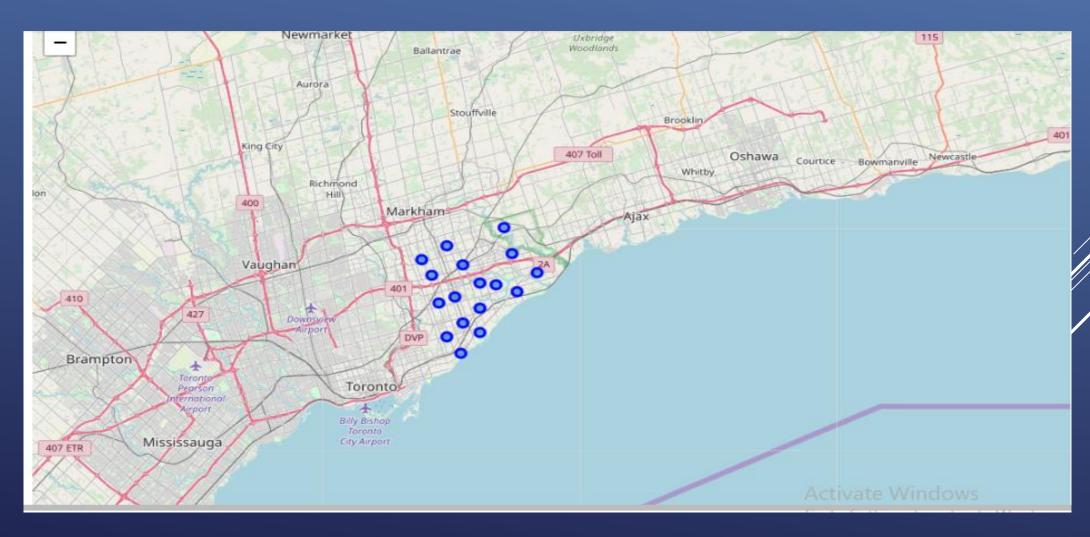
Data and methodology

- ▶ We have 103 observations with 5 features each in the final dataset
- ▶ As we concentrate on Scarborough only, we filter the data on this borough only.
- ▶ Based on the database, we collect data on 100 most common venues within 500 meters with the help of the Foursquare location data.
- ▶ We get 89 venues referring to one of the 53 categories.

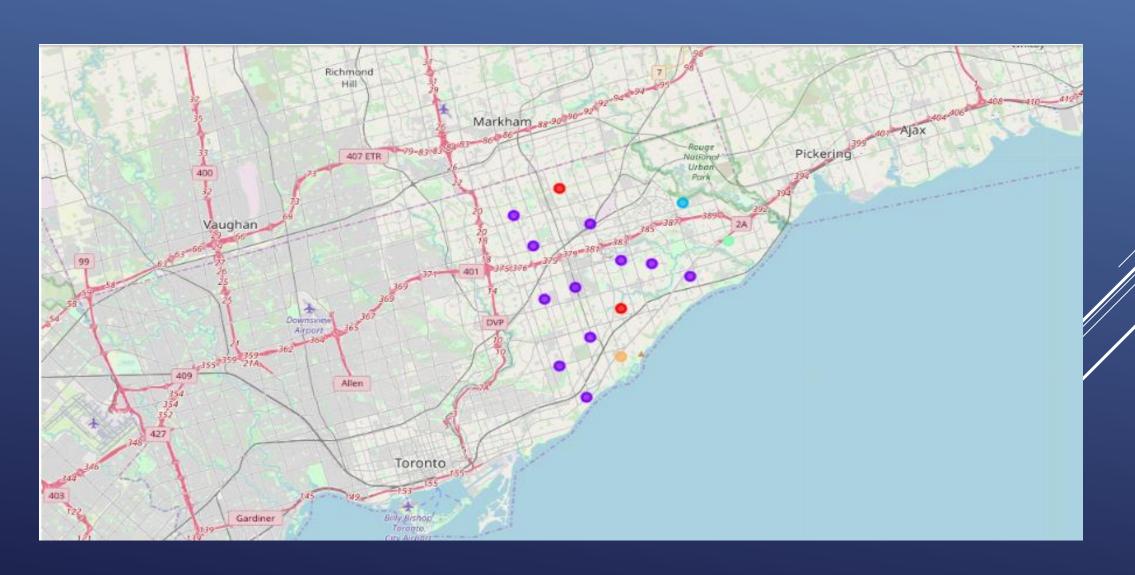
Neighborhoods of Toronto



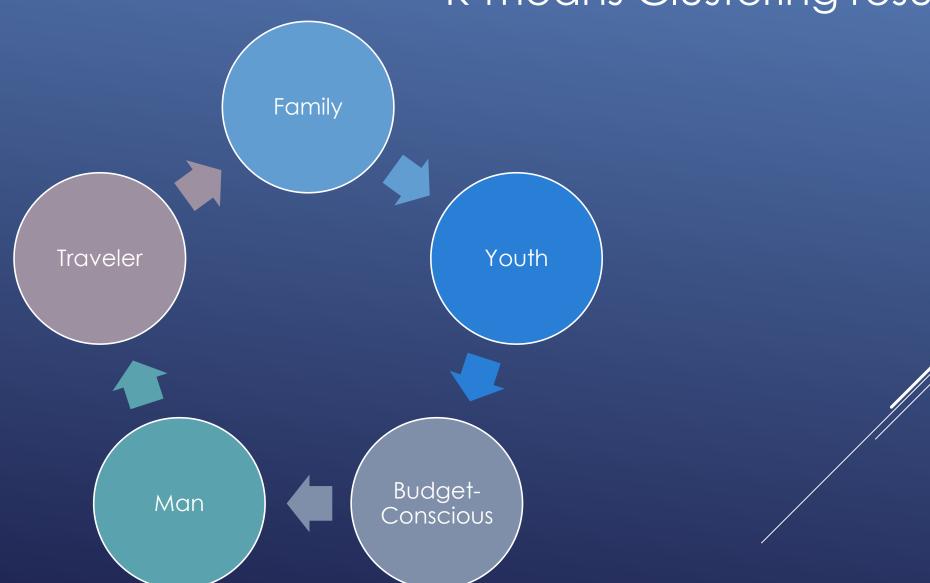
Neighborhoods of Scarborough, Toronto



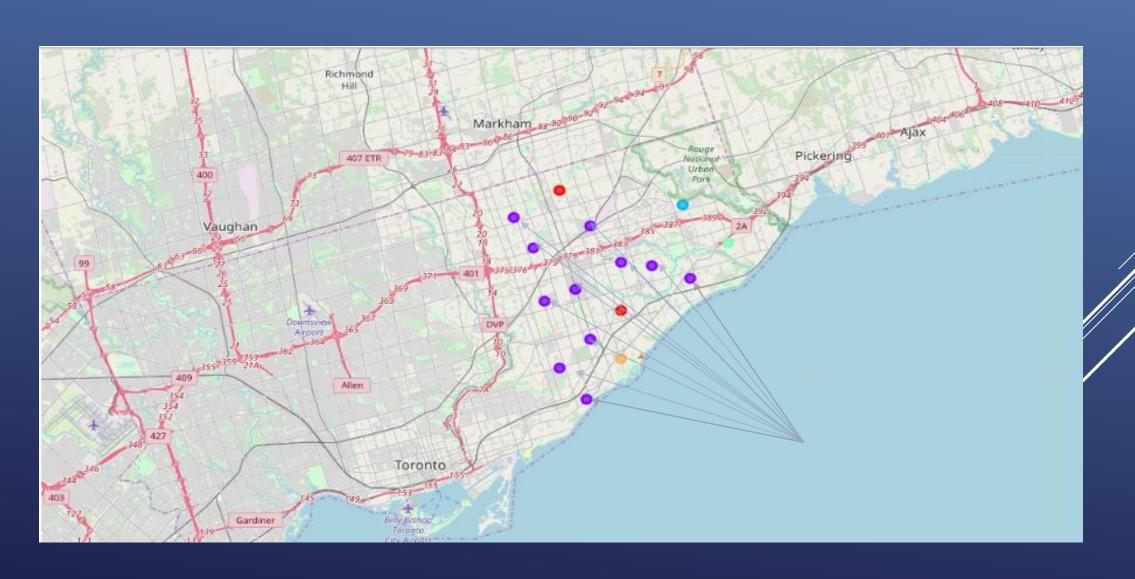
K-means Clustering results



K-means Clustering results



"Youth" cluster neighborhoods



Conclusion And Future Directions

- ► The analysis turned out really successful and hit all mentioned goals.
- Neighborhoods of the cluster "Youth" are recommended
 - Maximized demand here the common ones venues are mainly popular with the Youth (e.g. General Entertainment, Coffee Shops, Café's, Discount Stores or Fast Food Restaurants)
 - ▶ Minimized competition no clubs were mentioned.
- Make the same analysis within the chosen cluster only
- ▶ It might be useful to also analyze the data for the whole Toronto area.