

# CAPSTONE PROJECT- DATA SCIENCE

---

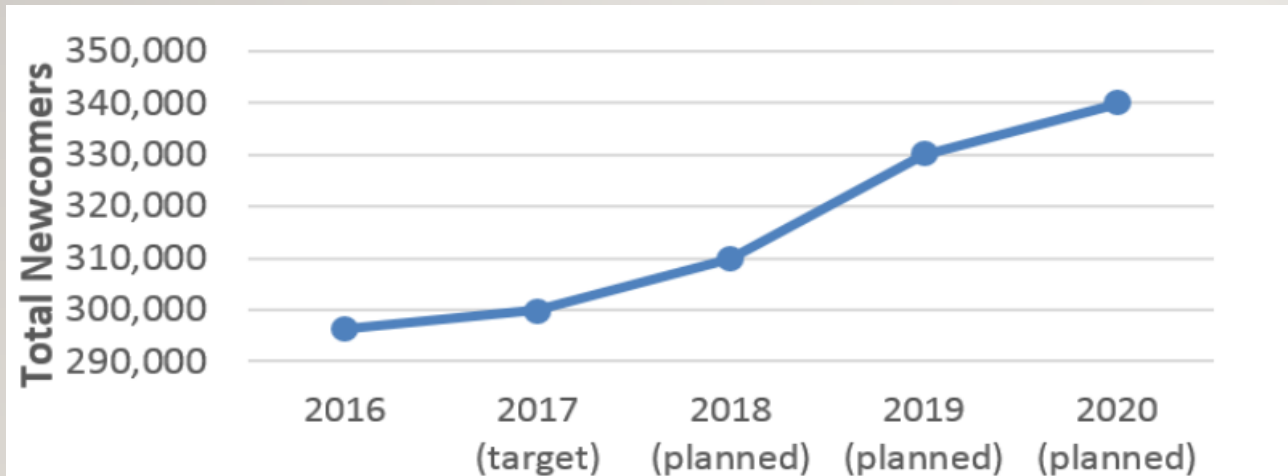
*USING SEGMENTING AND CLUSTERING TECHNIQUES TO OPEN A  
RESTAURANT IN TORONTO*

*PREPARED BY : GAVIN FERNANDES*



# BUSINESS CASE

- Immigrants to Canada has increased in the past year
- The total number of migrants is expected to increase by 340,000 by year 2020
- Locating a restaurant has always been a concern for newcomers coming with different taste and preferences
- The goal of this exercise would be to use a clustering & segmentation approach to select a borough in Toronto, being a capital city to open a restaurant



# METHODOLOGY

## Data Sourcing

- Collect Data from Wikipedia and source files from Toronto.

## Pre-Process Data

- Use Webscraping Techiques to extract data from Wikipedia.
- Impute Data into pandas dataframe.
- Remove null values of missing values.

## Enhance Data using API

- Use Foursquare API to import nearby Venues Categories. using Latitude and Longitude of Toronto.

## Clustering and Segmentation

- Use K means clustering algorithm to group venues categories in Toronto with similar characteristics.

## Select Location

- Select the location which has the cluster with the maximum count.

# CLUSTERING ALGORITHM - BEFORE CLUSTERING

	PostalCode	Borough	Latitude	Longitude	Places	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
0	M4E	East Toronto	43.676357	-79.293031	The Beaches	Neighborhood	Pub	Coffee Shop	Health Food Store
1	M4K	East Toronto	43.679557	-79.352188	The Danforth West, Riverdale	Greek Restaurant	Coffee Shop	Italian Restaurant	Ice Cream Shop
2	M4L	East Toronto	43.668999	-79.315572	The Beaches West, India Bazaar	Pet Store	Ice Cream Shop	Pizza Place	Movie Theater
3	M4M	East Toronto	43.659526	-79.340923	Studio District	Café	Coffee Shop	Bakery	Italian Restaurant



# CLUSTERING ALGORITHM - AFTER CLUSTERING

	PostalCode	Borough	Latitude	Longitude	Places	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
0	M4E	East Toronto	43.676357	-79.293031	The Beaches	4	Neighborhood	Pub	Coffee Shop	Health Food Store
1	M4K	East Toronto	43.679557	-79.352188	The Danforth West, Riverdale	0	Greek Restaurant	Coffee Shop	Italian Restaurant	Ice Cream Shop
2	M4L	East Toronto	43.668999	-79.315572	The Beaches West, India Bazaar	0	Pet Store	Ice Cream Shop	Pizza Place	Movie Theater
3	M4M	East Toronto	43.659526	-79.340923	Studio District	0	Café	Coffee Shop	Bakery	Italian Restaurant

# CLUSTERING ALGORITHM - CONCLUSION

```
0    33
1     2
4     1
3     1
2     1
Name: Cluster Labels,
```

***It can be concluded that since the majority of places with similar venue categories are surrounded in the East Toronto Area  
It would be prudent to open a restaurant in this area as a business.***