

APPLIED DATA SCIENCE CAPSTONE PROJECT

Recommend location for a new Restaurant through
Data Analysis

Nithin John Jacob

RECOMMENDING LOCATION FOR A NEW INDIAN RESTAURANT IN TORONTO

- Setting up a new restaurant –Huge investment
- Location, Location & Location
- Recommend location based on demographics and restaurant profile

DATA

- **Toronto Demographics Data**

<https://open.toronto.ca/dataset/neighbourhood-profiles/>.

Population Area, Average income Ethnicity ...etc

- **Geographic data of Toronto neighbourhoods**

<https://github.com/jasonicarter/toronto-geojson>

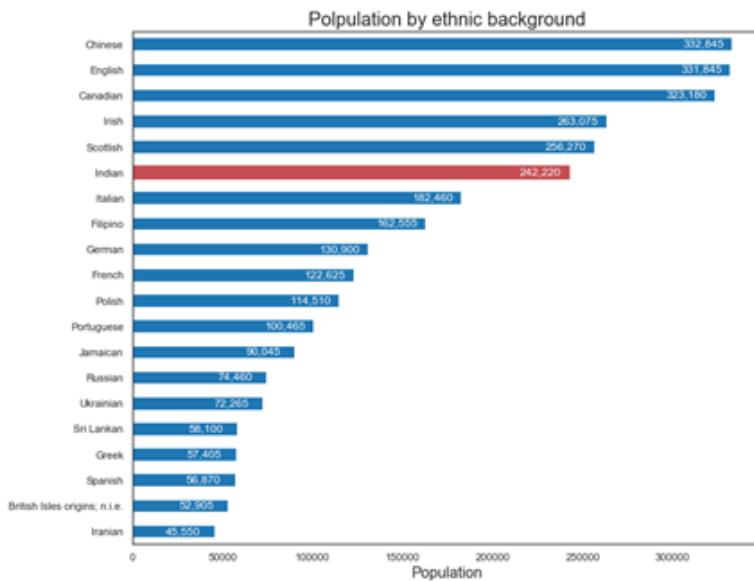
The boundaries of each neighbourhood >> centroid

- **Restaurant Data from Foursquare**

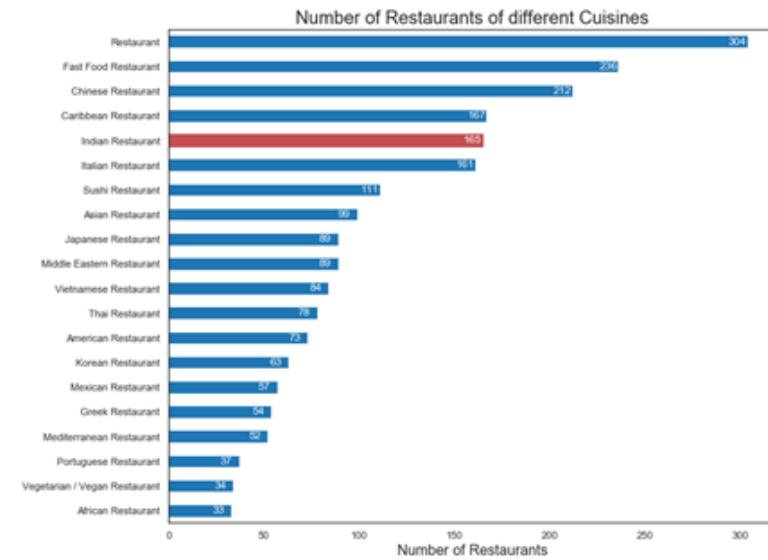
Restaurants, lat, long, Type, Address

EXPLORATORY DATA ANALYSIS

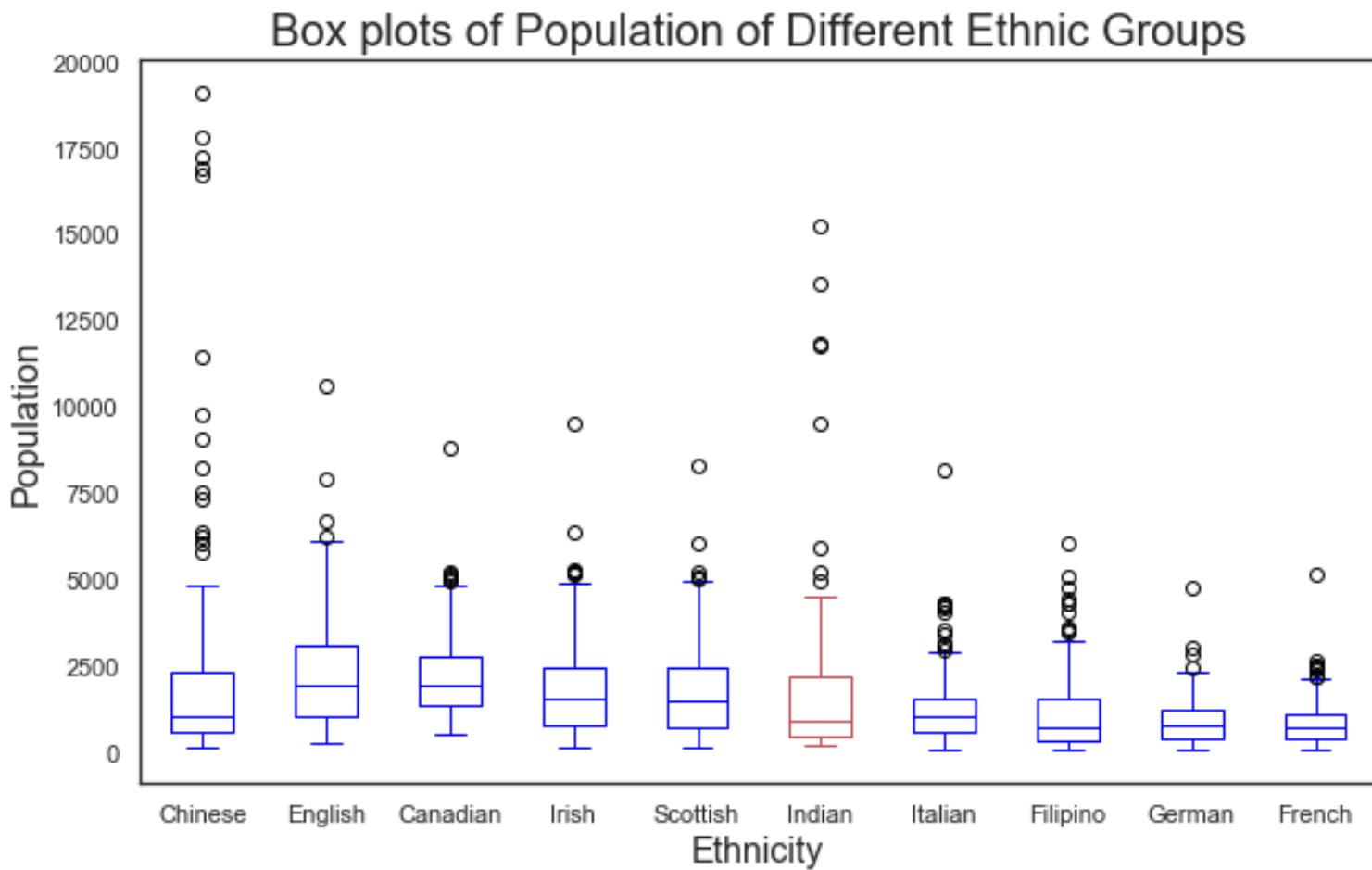
Population



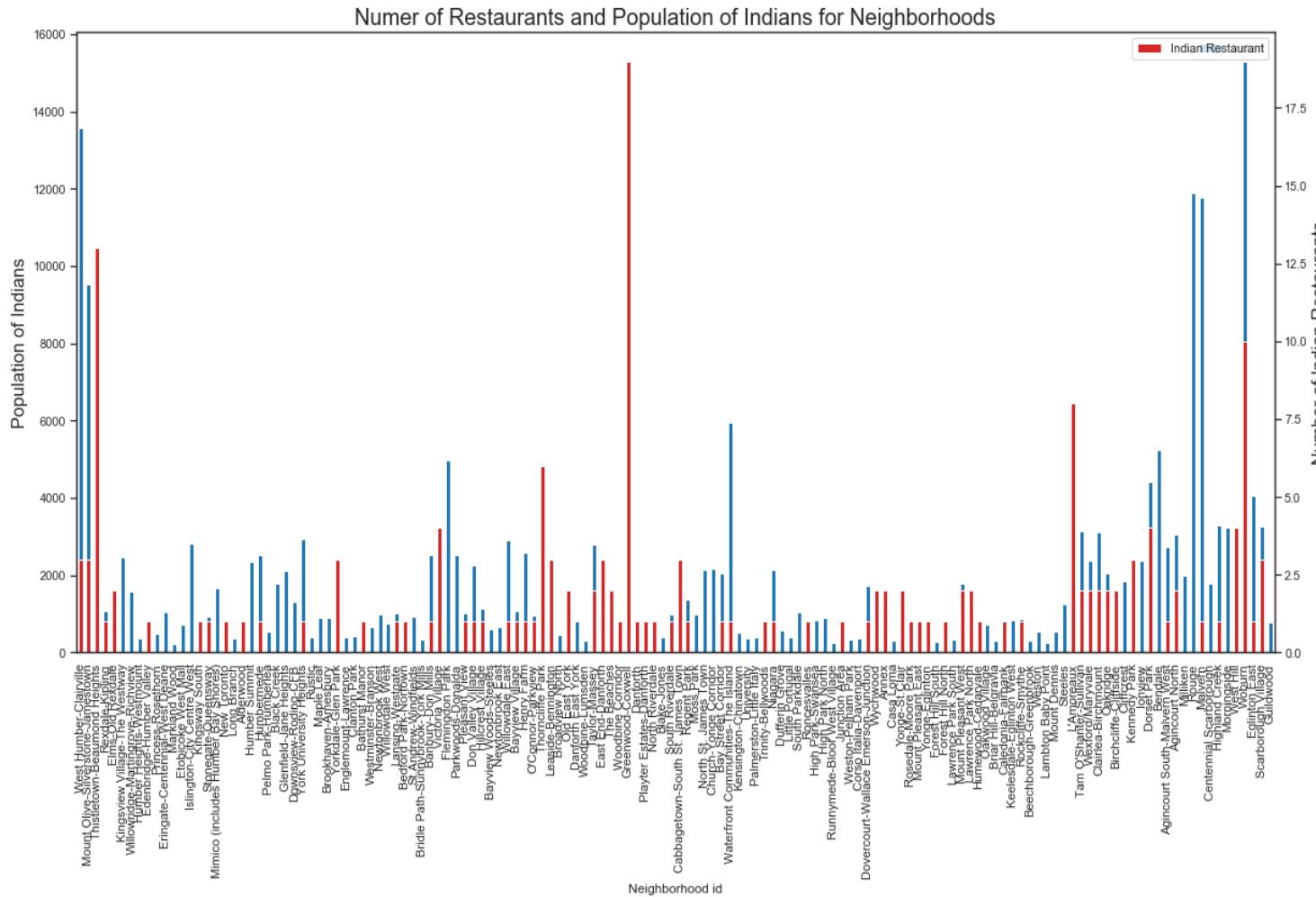
Restaurants



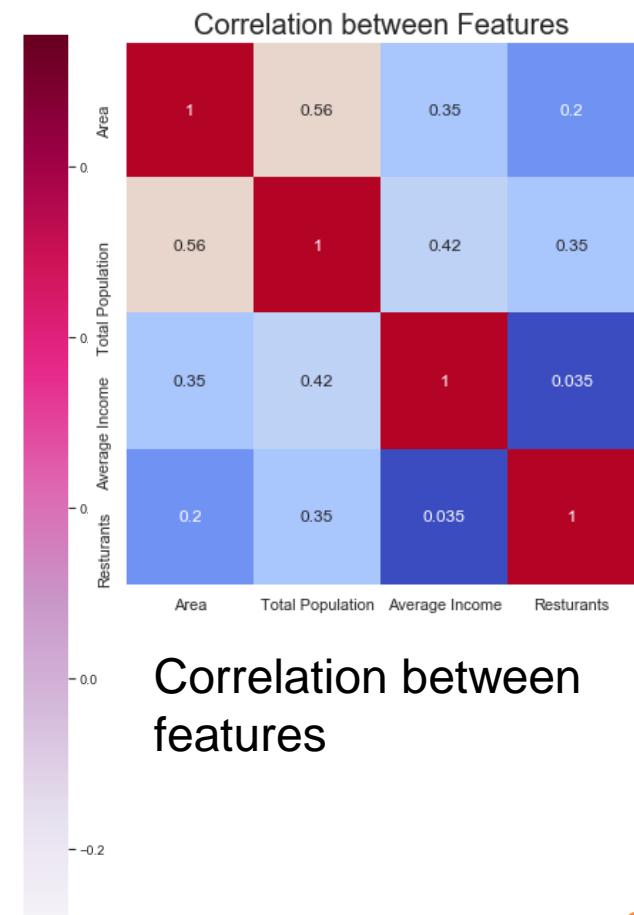
POPULATION DISTRIBUTION



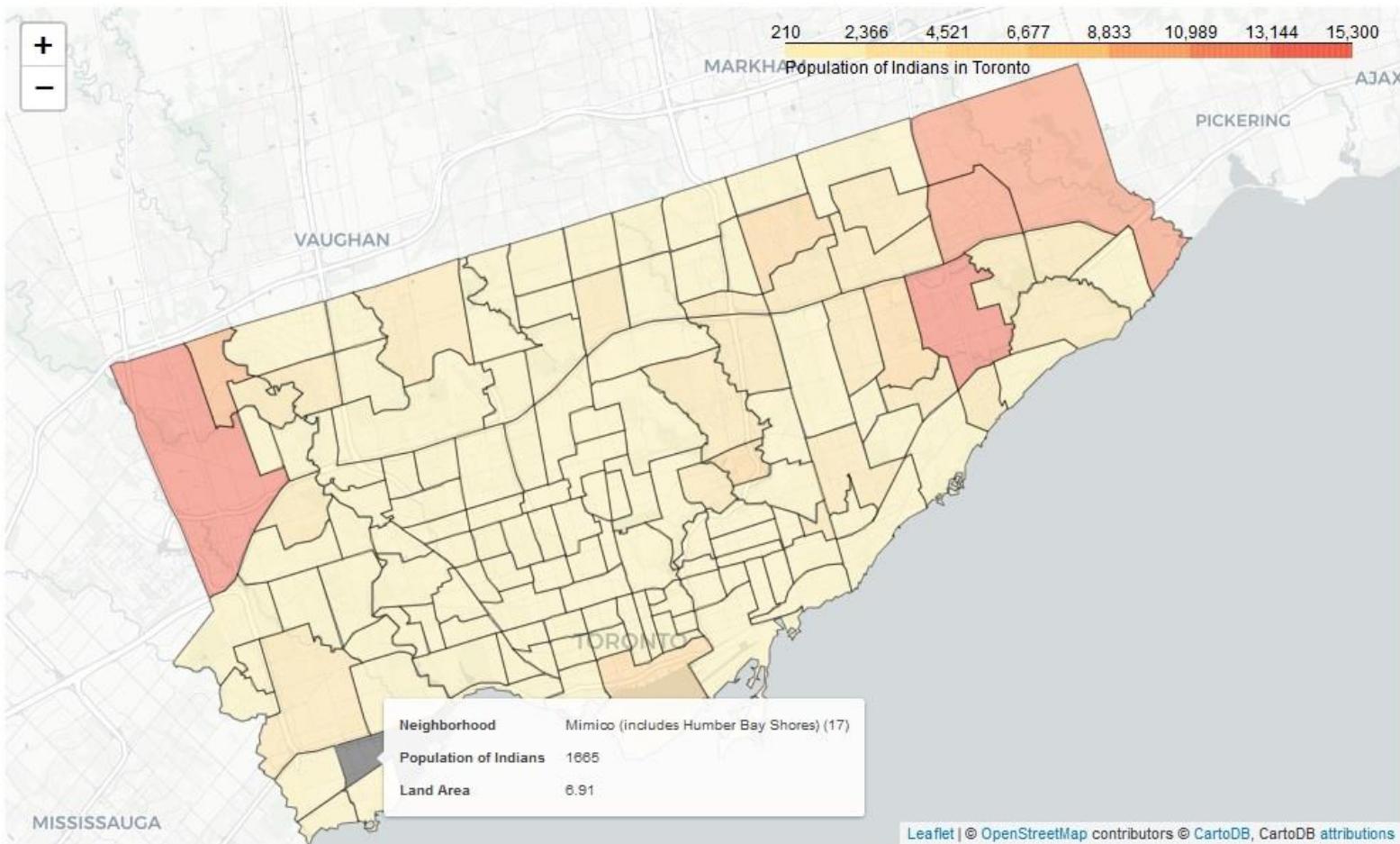
INDIAN POPULATION AND RESTAURANTS



CORRELATION

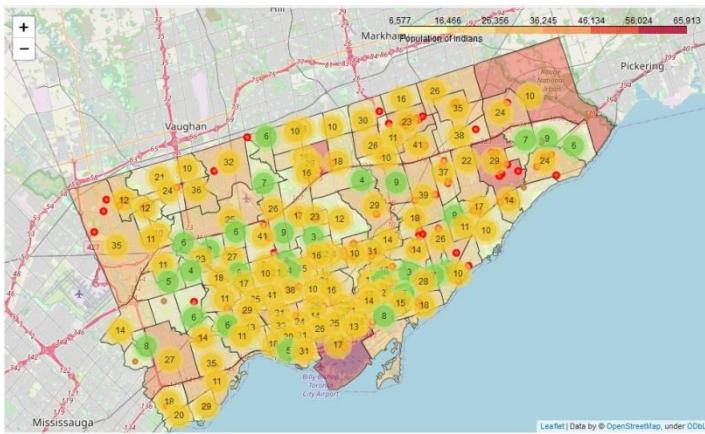


HEAT MAP OF INDIANS

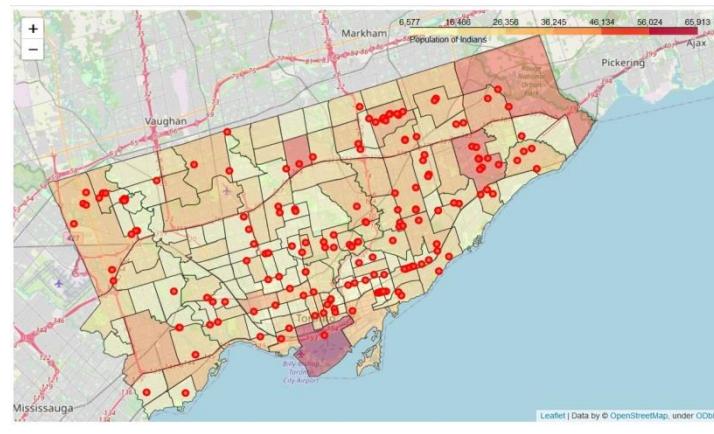


DATA IN MAP

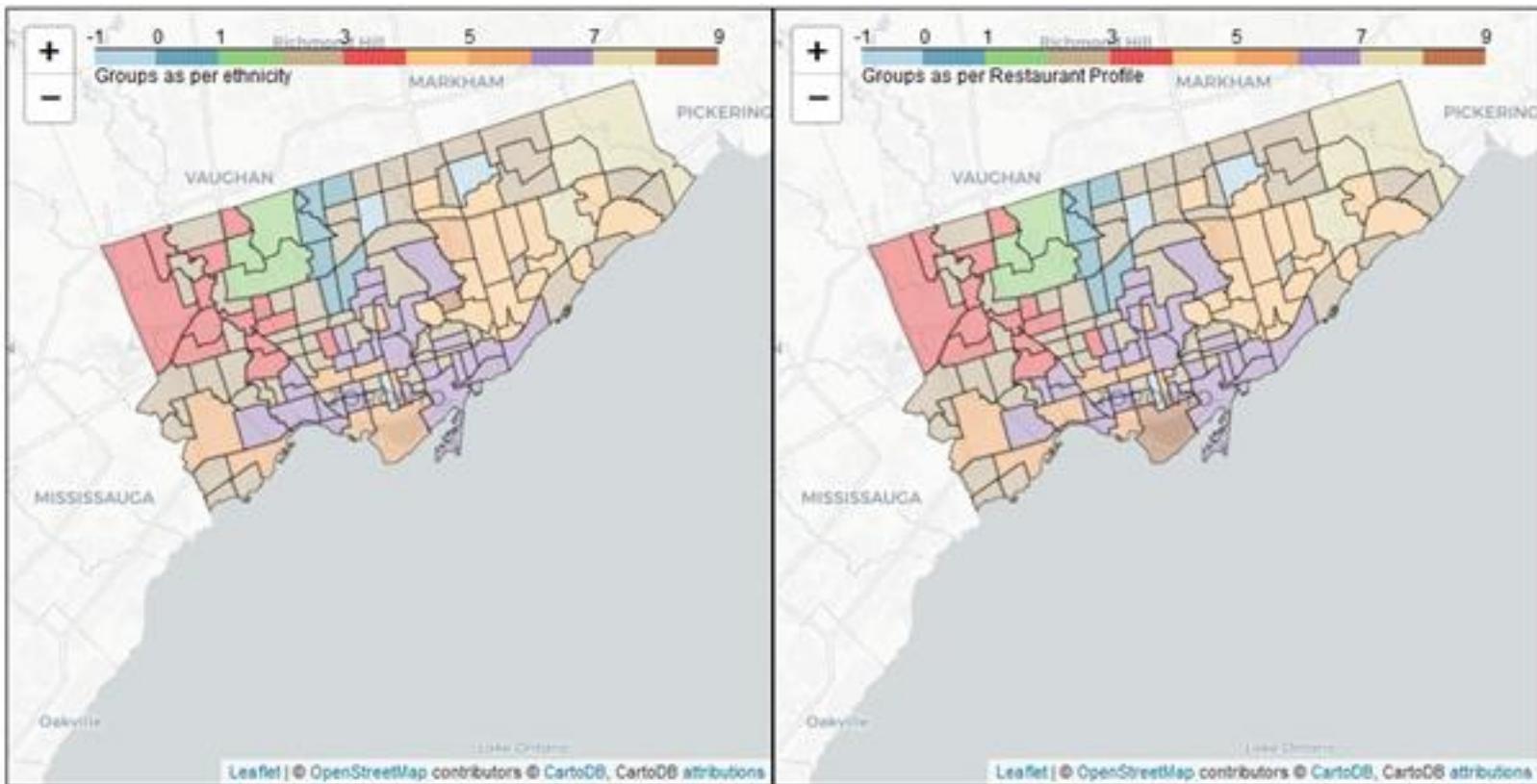
- All Restaurants



- Indian Restaurants



K-MEANS CLUSTERING



By Demographics

By Restaurant Profiles

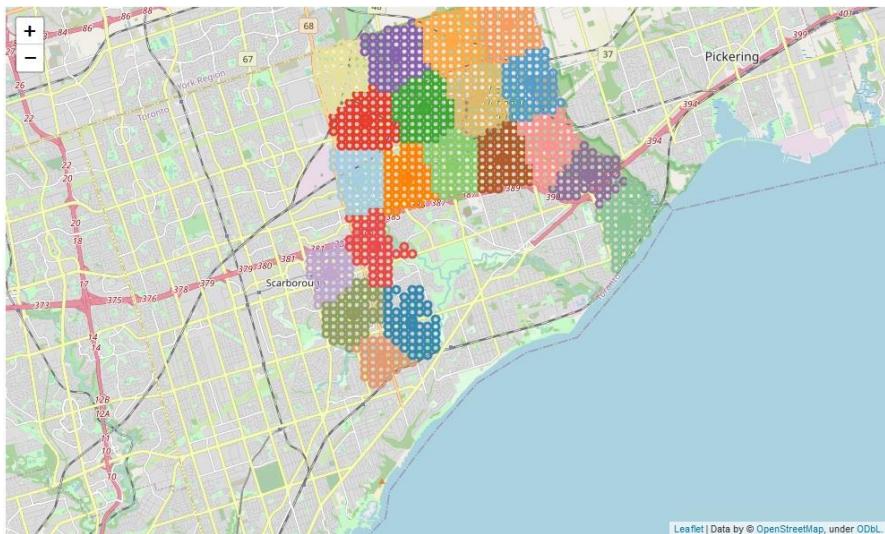
RESULTS OF K-MEANS

- Two cluster groups sorted in descending order pf population of Indians
- Woburn is selected as it has maximum Population among top 3

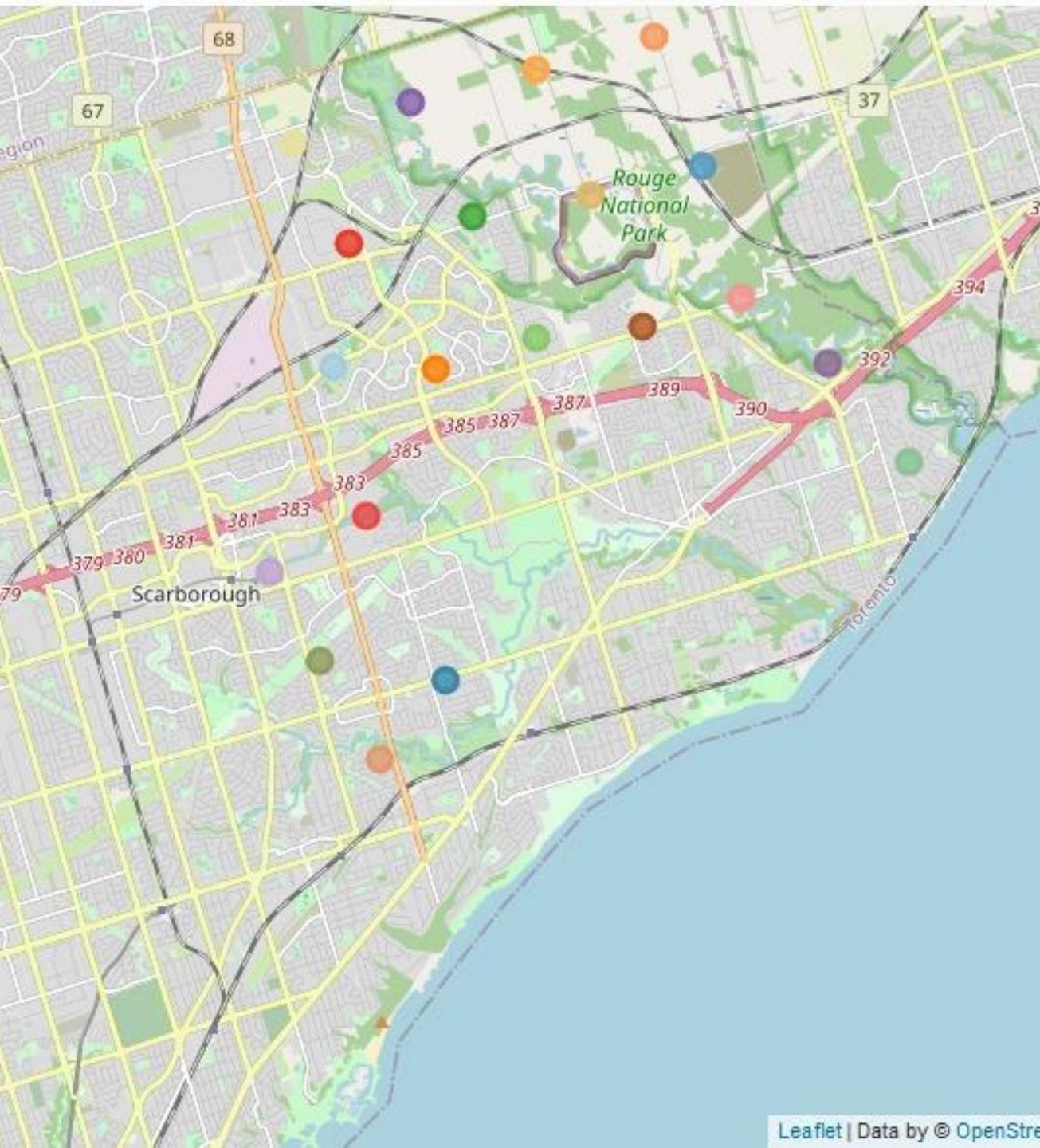
Neighborhood id	Neighborhood	Indian Restaurant	Restaurant Cluster	Demographics Cluster	Indian Population
65	Greenwood-Coxwell	19	3	3	725
3	Thistletown-Beaumont Heights	13	3	1	2485
137	Woburn	10	9	9	16390
117	L'Amoreaux	8	1	0	4765
55	Thorncliffe Park	6	1	0	5020
136	West Hill	4	1	5	3480
126	Dorset Park	4	1	1	4800
43	Victoria Village	4	1	0	2230
1	West Humber-Clairville	3	5	5	14125

SELECTED NEIGHBORHOODS

Neighborhood id	Neighborhood	Indian
131	Rouge	12580
132	Malvern	12640
137	Woburn	16390



- Selected cluster with Woburn
- Grid 200x200m in the selected Neighborhoods
- Remove points near indian restaurant
- Group remaining points to 20 groups
- Centroid of location as recommended places



RECOMMENDED LOCATIONS

The centroids of cluster groups are identified as the recommended locations .

The addressess are obtained using geopy

Address1:
103, Littleleaf Drive, Browns
Corners, Scarborough North,
Scarborough, Toronto, Golden
Horseshoe, Ontario, M1B 1P6,
Canada

CONCLUSION

- The Data was collected and analysed using statistical and machine learning techniques and twenty locations are identified for further exploration.