CAPSTONE PROJECT:
BATTLE OF NEIGHBORHOODS

-- Suresh

PROJECT DESCRIPTION:

- Analysis of neighborhoods in Scarborough (Toronto) based on the distribution of various facilities in and around the neighborhood to help the new Neighbors to take a better decision on choosing the best neighborhood out of Scarborough.

OBJECTIVE :

- The main purpose of this project is to help new neighbors to explore the various neighborhoods before actually move in to any neighborhood.

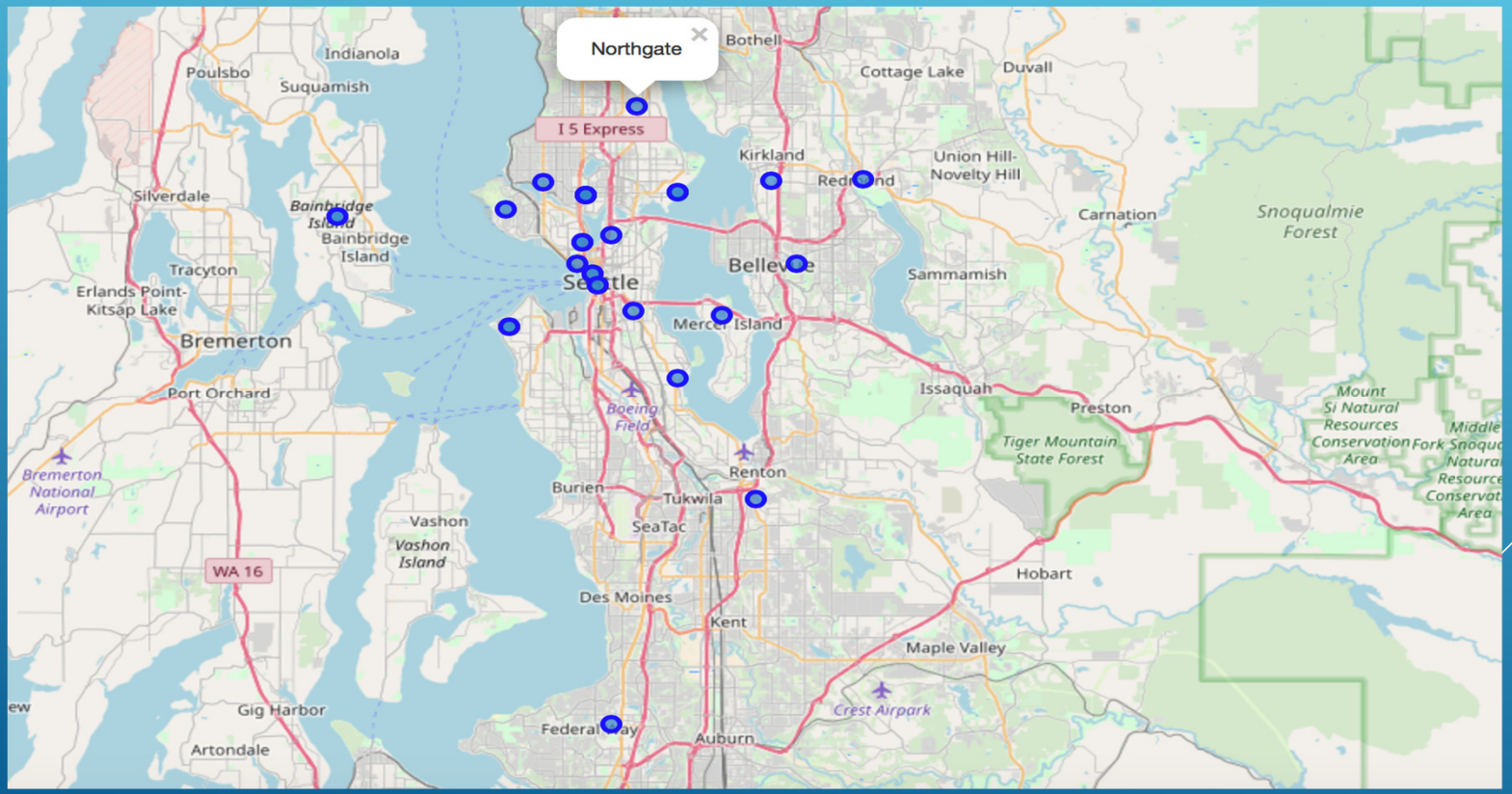
DATA AND METHODOLOGY :

- Wikipedia page, https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M, has been used in order to obtain the neighborhood data along with postal codes and transformed the data into a pandas dataframe by using BeatifulSoup Pacakge
- Forming neighborhood clusters based on venue categories using unsupervised k-mean clustering algorithm(sklearn)
- Identifying and understanding the similarities and dissimilarities between two chosen neighborhoods to retrieve more insights and to conclude with ease which neighborhood wins over other.

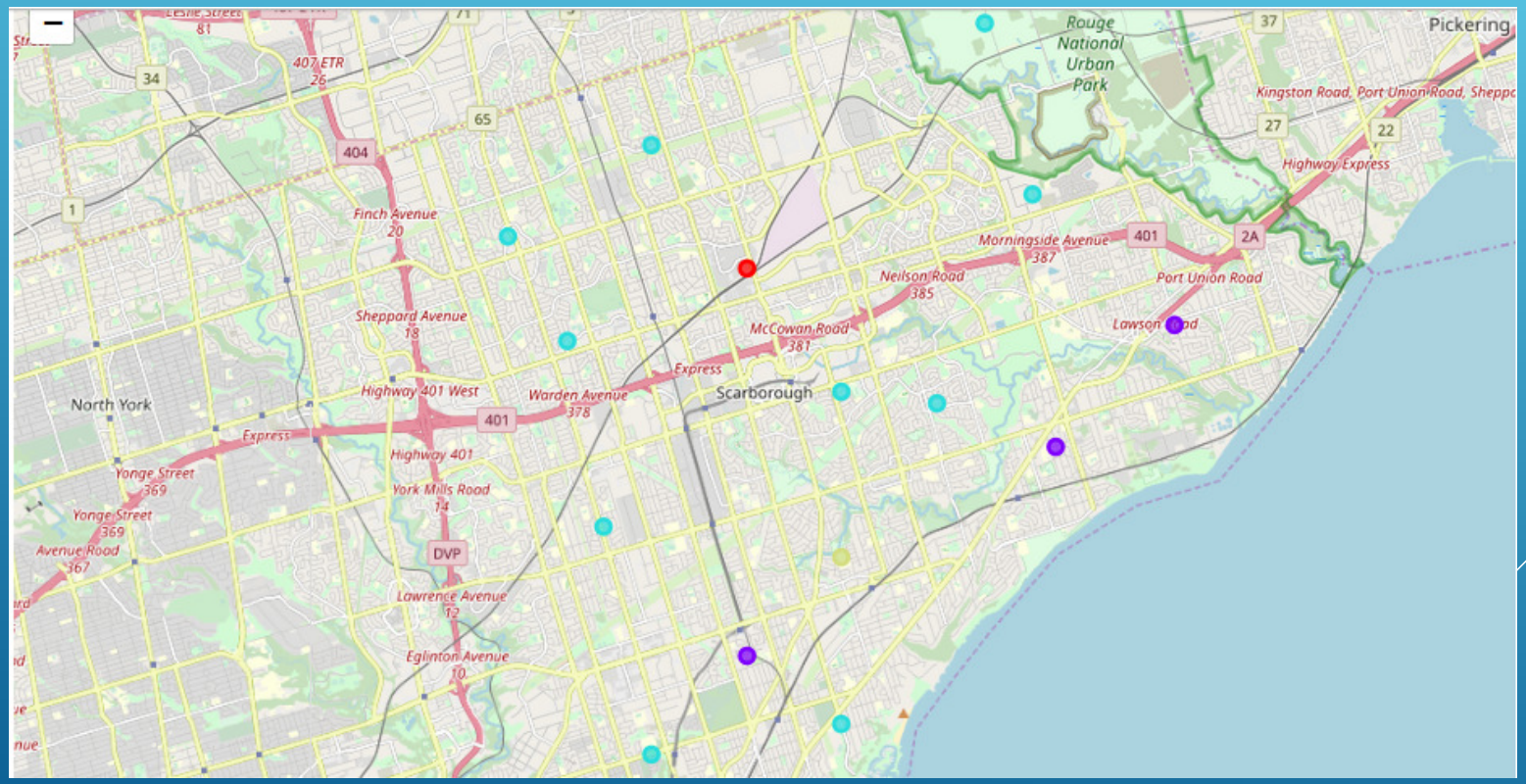
PYTHON PACKAGES :

- Pandas - Library for Data Analysis
- NumPy – Library to handle data in a vectorized manner
- JSON – Library to handle JSON files
- Geopy – To retrieve Location Data
- Requests – Library to handle http requests
- Matplotlib – Python Plotting Module
- Sklearn – Python machine learning Library
- Folium – Map rendering Library

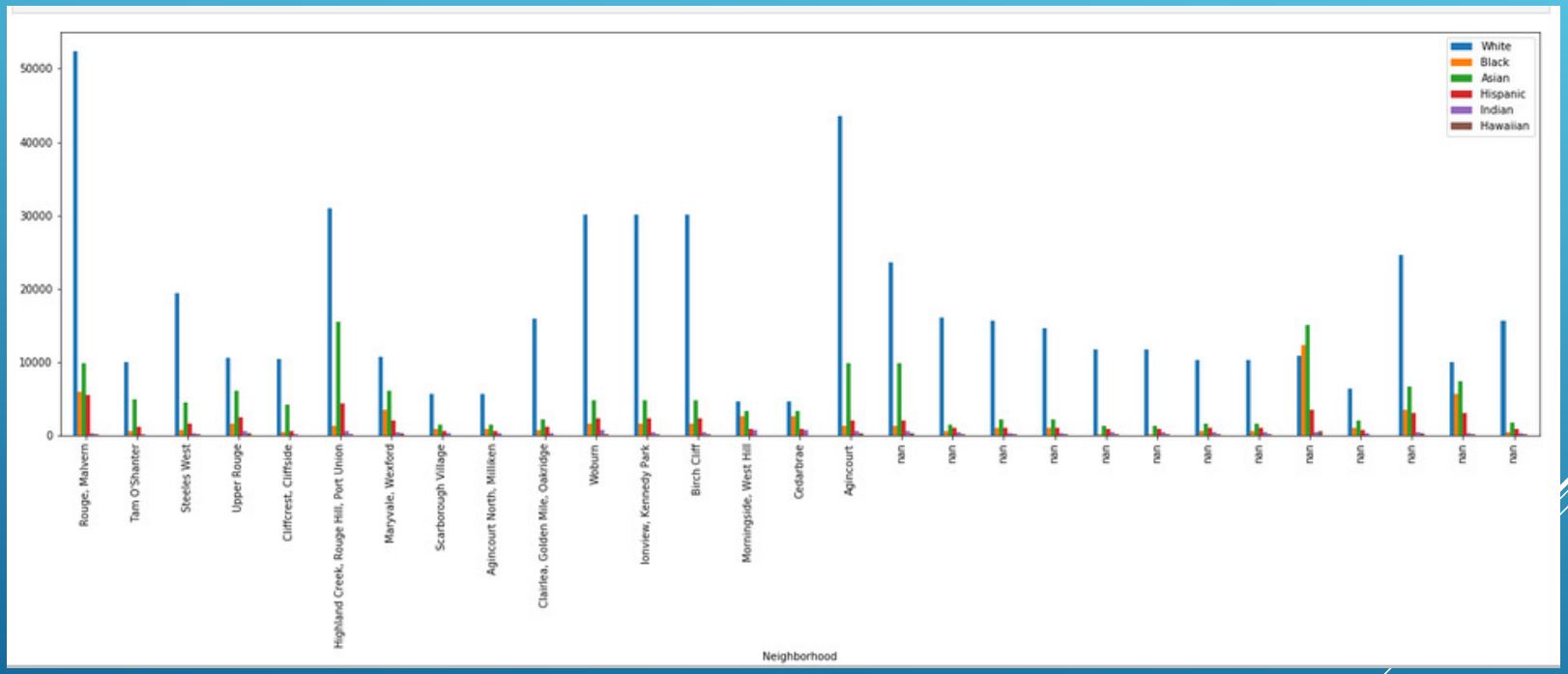
NEIGHBORHOOD DISTRIBUTION – FOLIUM MAP



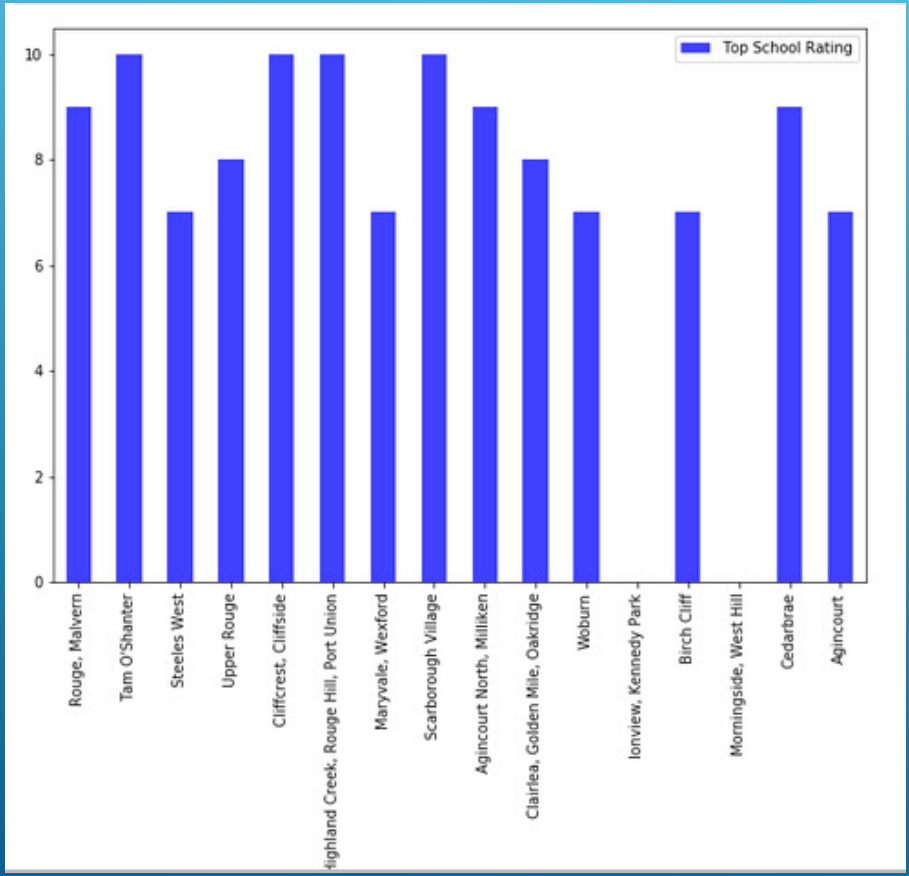
NEIGHBORHOOD CLUSTERS :



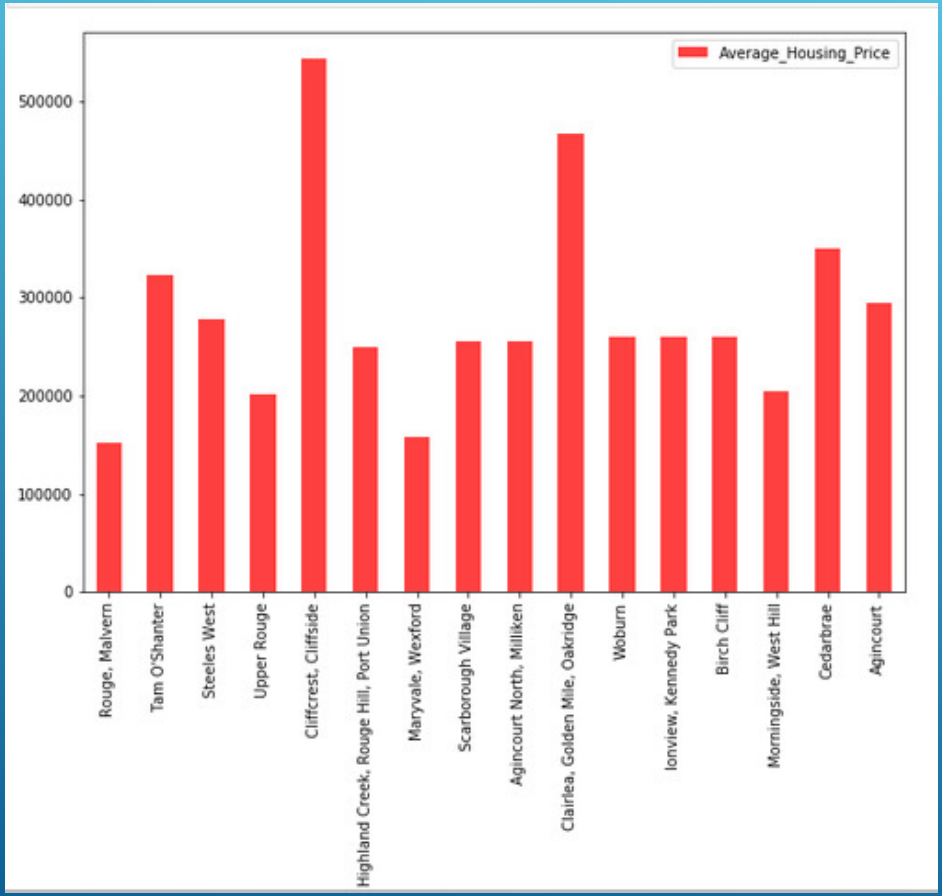
POPULATION DISTRIBUTION :



SCHOOL RATINGS :



AVERAGE HOUSING PRICES :



COMPARISON BETWEEN NEIGHBORHOODS :

K-mean clustering unsupervised machine learning algorithm has been used to cluster the neighborhoods.

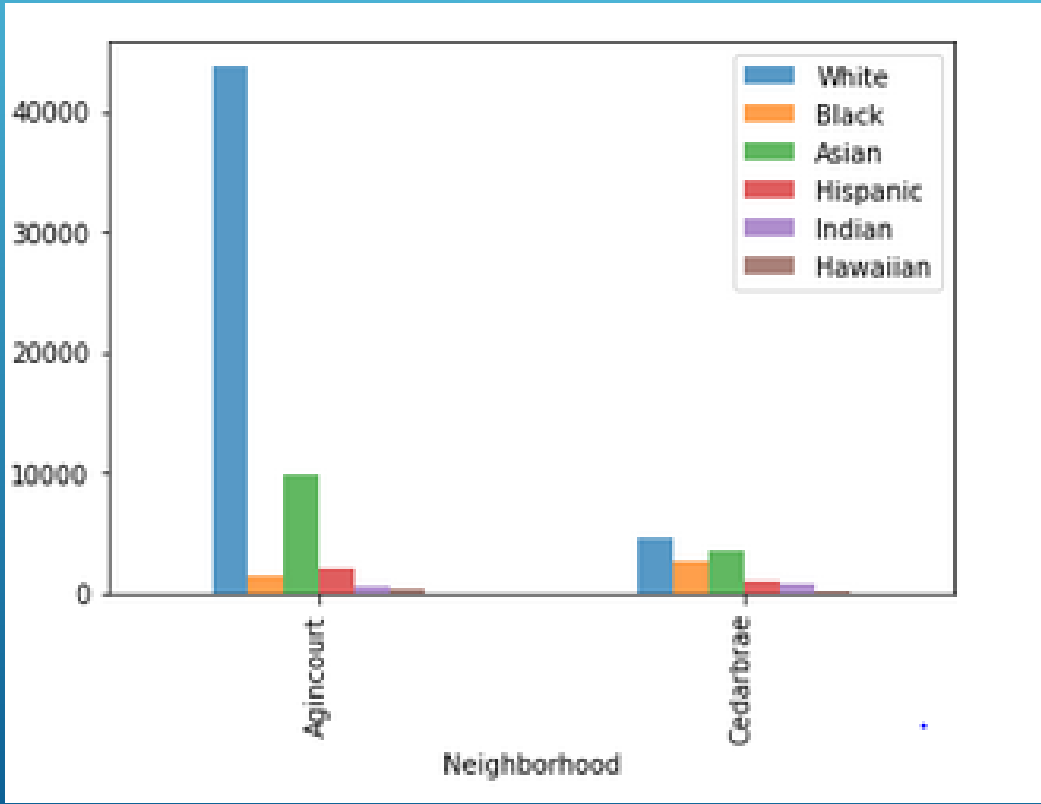
Agincourt, Cedarbraeto are compared on the following parameters*:

1. More Indian Population
2. Higher School Rating
3. Reasonable Housing Price in the Range of 150k to 500k
4. Comfortable Neighborhoods.

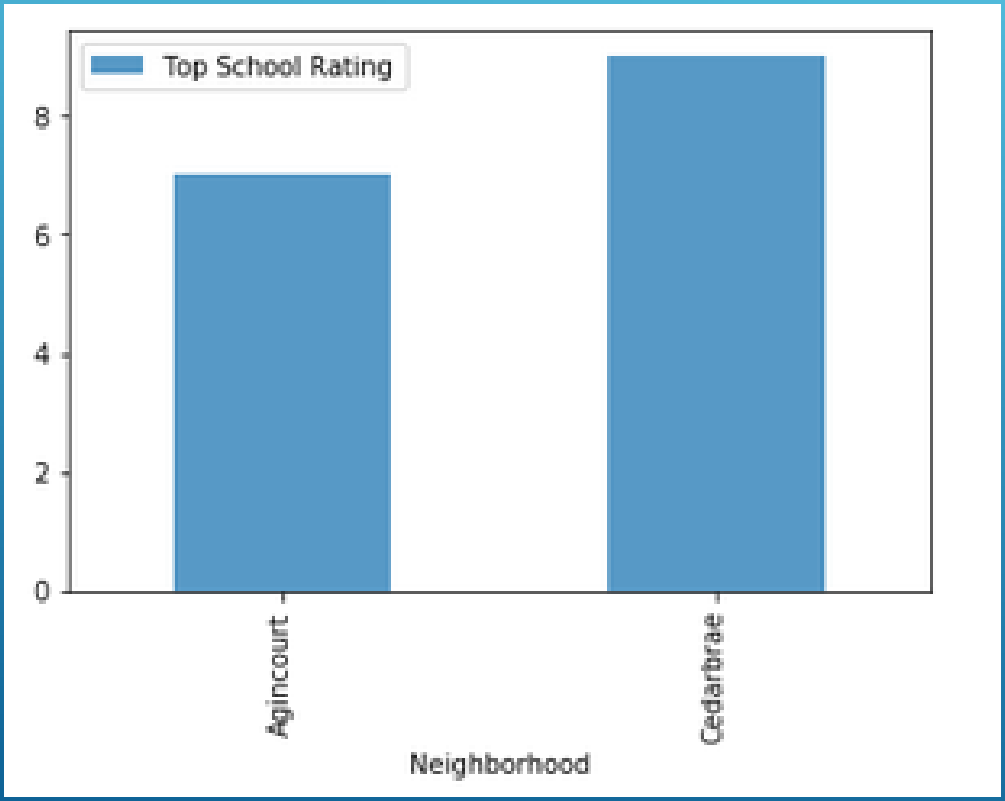
NEIGHBORHOOD VENUES:

Neighborhood	Agincourt	Cedarbrae
Postcode	M1S	M1H
Borough	Scarborough	Scarborough
Latitude	43.7942	43.7731
Longitude	-79.262	-79.2395
Cluster Labels	0	2
1st Most Common Venue	Clothing Store	Indian Restaurant
2nd Most Common Venue	Pool Hall	Coffee Shop
3rd Most Common Venue	Motorcycle Shop	Bakery
4th Most Common Venue	Badminton Court	Hakka Restaurant
5th Most Common Venue	Lounge	Lounge
6th Most Common Venue	Skating Rink	Gym / Fitness Center
7th Most Common Venue	Shanghai Restaurant	Chinese Restaurant
8th Most Common Venue	Breakfast Spot	Caribbean Restaurant
9th Most Common Venue	Construction & Landscaping	Flower Shop
10th Most Common Venue	Convenience Store	Rental Car Lo

ETHNIC POPULATION DISTRIBUTION :



SCHOOL RATINGS:



AVERAGE HOUSING PRICE :

DISCUSSION AND CONCLUSION :

- This Analysis concludes that compared to Agincourt, Cedarbrae has the higher number of Indian population, good school rating of 9 and a reasonable avg housing price of around 360K ,also top 10 common venues shows Cedarbrae has got a good neighborhood with Indian Restaurant, Cafe Shop, Bakery Shop, Hakka Restaurant etc.

- Hence Cedarbrae wins over Agincourt!

Thank You