<u>Capstone Project-Battle of Neighborhoods:</u>

Is Ahmedabad City similar to Toronto city?

Mohit Rajput

Table of Contents

Introduction

Ahmedabad is located in Gujarat, India. It plays a major role in Country's economy. Ahmedabad is known for his business capital and Restaurant businesses. Toronto is the capital city of the province of Ontario and the largest city in Canada. Toronto is a Centre of business, finance, arts, and culture, and is recognized as one of the most multicultural and cosmopolitan cities in the world. Both cities are well known in Malaysia, and become the top choice for local and foreign communities.

Brief information about both cities:

AHMEDABAD: Ahmadabad has emerged as an important economic and industrial hub in India. It is the second largest producer of cotton in India, and its stock exchange is the country's second oldest. Cricket is a popular sport in Ahmadabad, which houses the 54,000-seat Sardar Patel Stadium. The effects of liberalization of the Indian economy have energized the city's economy towards tertiary sector activities such as commerce, communication and construction. Ahmadabad's increasing population has resulted in an increase in the construction and housing industries resulting in recent development of skyscrapers.

(source: https://en.wikipedia.org/wiki/Ahmedabad)

TORONTO: Toronto is a prominent Centre for music, theatre, motion picture production, and television production, and is home to the headquarters of Canada's major national broadcast networks and media outlets. Its varied cultural institutions, which include numerous museums and galleries, festivals and public events, entertainment districts, national historic sites, and

sports activities, attract over 25 million tourists each year. Toronto is known for its many

skyscrapers and high-rise buildings, in particular the tallest free-standing structure in the

Western Hemisphere, the CN Tower.

(source: https://en.wikipedia.org/wiki/Toronto)

Objective (Business Problem)

In this project, we will study in details the area and doing classification using Foursquare data and machine learning segmentation and clustering. The aim of this project is to segment areas of Ahmedabad and Toronto based on the most common places captured from Foursquare.

Using segmentation and clustering, we hope we can determine:

- The similarity or dissimilarity of both cities
- Classification of area located inside the city whether it is residential, tourism places, or others
- Find an optimal location for establish a business
- Restaurant food popularity
- Identify optimal place for Bus station, Railway station or Airport

Data

The data we used in this project is in .csv format. Hence, there are no accurate data is available for Ahmedabad City, so all data is collected by manually and Wikipedia. For, Toronto City we used Wikipedia's List_of_postal_codes_of_Canada__M.html and Toronto geospatial data.

The data acquired from Wikipedia pages and restructure to csv file for easier manipulation and reading. Both Ahmedabad and Toronto city necessary data files uploaded to my GitHub link for references. In, this data each row is referred by neighborhood.

Links to websites which are used to get data:

- `https://en.wikipedia.org/wiki/Category:Neighbourhoods_in_Ahm edabad`

- `https://www.latlong.net/`

- `https://www.census2011.co.in/census/city/314-ahmedabad.html`

Link to the data files are:

- `https://github.com/mohit-n-rajput/Coursera_Capstone/blob/master/ data/Geospatial_Coordinates.csv`

- `https://github.com/mohit-n-rajput/Coursera_Capstone/blob/master/ data/List_of_postal_codes_of_Canada__M.html`

- `https://github.com/mohit-n-rajput/Coursera_Capstone/blob/master/ data/Geospatial_Coordinates.csv`

Another aspect to consider for this project is the Foursquare data. I believe that the data as good as provided, meaning although we are using Foursquare data for segmentation and clustering, the amount and accuracy of data captured can't 100% determine correct classification in real world. Also, there are no accurate data about Ahmedabad neighborhood, it's all collected by me.

Ahmedabad City Neighborhood Data:

```
In [3]: #create dataframe
        df_to = pd.read_csv('ahmedabad.csv',delimiter=',', header = None)
        df_to.columns = ['Neighbourhood','Latitude','Longitude']

        df_to.head(10)
```

Out[3]:

	Neighbourhood	Latitude	Longitude
0	Amraiwadi	23.010590	72.619034
1	Asarwa	23.044980	72.607700
2	Ashram Road	23.006986	72.557490
3	Astodia	23.016830	72.590975
4	Bapunagar	23.032660	72.629494
5	Behrampura	23.003589	72.583840
6	Bodakdev	23.044371	72.517921
7	Bopal	23.030050	72.464943
8	CG Road	23.039084	72.562786
9	Chandkheda	23.112650	72.583618

Toronto City Neighborhood Data:

```
In [169]:   df_neighbours = pd.merge(df_to,toronto_geo_data,on=['Postcode'],how='inner')
            df_tor.head(15)
```

Out[169]:

	Postcode	Borough	Neighbourhood	Latitude	Longitude
0	M1B	Scarborough	Rouge, Malvern	43.806686	-79.194353
1	M1C	Scarborough	Highland Creek, Rouge Hill, Port Union	43.784535	-79.160497
2	M1E	Scarborough	Guildwood, Morningside, West Hill	43.763573	-79.188711
3	M1G	Scarborough	Woburn	43.770992	-79.216917
4	M1H	Scarborough	Cedarbrae	43.773136	-79.239476
5	M1J	Scarborough	Scarborough Village	43.744734	-79.239476
6	M1K	Scarborough	East Birchmount Park, Ionview, Kennedy Park	43.727929	-79.262029
7	M1L	Scarborough	Clairlea, Golden Mile, Oakridge	43.711112	-79.284577
8	M1M	Scarborough	Cliffcrest, Cliffside, Scarborough Village West	43.716316	-79.239476
9	M1N	Scarborough	Birch Cliff, Cliffside West	43.692657	-79.264848
10	M1P	Scarborough	Dorset Park, Scarborough Town Centre, Wexford ...	43.757410	-79.273304
11	M1R	Scarborough	Maryvale, Wexford	43.750072	-79.295849
12	M1S	Scarborough	Agincourt	43.794200	-79.262029
13	M1T	Scarborough	Clarks Corners, Sullivan, Tam O'Shanter	43.781638	-79.304302
14	M1V	Scarborough	Agincourt North, L'Amoreaux East, Milliken, St...	43.815252	-79.284577

City data accessible through the FourSquare API:

This data will produce the following deliverables.

- A list of all neighborhoods in Toronto, Ahmedabad with a different venues

- A list of the top ten venues in each neighborhood

- A ranking of all neighborhoods with a specific venue based on Review

Methodology

In this project, we will use the basic methodology of data science methodology.

After the Business Problem Definition, Best analytic approach is chosen.

After that process of data analysis is done:

- Data requirements

- Data collection

- Data processing

- Data cleaning

- Exploratory data analysis

- Modeling and algorithms

- Data product

Data requirements, collection, preprocessing:

In this step data about like neighborhood, and location of neighbor of Ahmedabad and Toronto Cities are collected, after that NaN values are removed and data is preprocessed.

After Data Preprocessing:

- We have done convert addresses into their equivalent latitude and longitude values, and neighborhoods name.

- Then we will use the Foursquare API to explore neighborhoods in both cities, Ahmedabad and html

- After that, explore function to get the most common venue categories in each neighborhood, and then use this feature to group the neighborhoods into clusters

Modeling and algorithms:

K-means clustering algorithm will be use to complete this task. And also, **the Folium library to visualize the neighborhoods** in Ahmedabad and Toronto and their emerging clusters.

Here K-Means Clustering Intuition are Shown:

K-MEANS CLUSTERING

- ## Description

Given a set of observations (x_1, x_2, …, x_n), where each observation is a d-dimensional real vector, k-means clustering aims to partition the n observations into k sets ($k \leq n$) $\mathbf{S} = \{S_1, S_2, …, S_k\}$ so as to minimize the within-cluster sum of squares (WCSS):
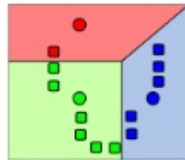
$$\arg\min_{S} \sum_{i=1}^{k} \sum_{x_j \in S_i} \left\| x_j - u_i \right\|^2$$
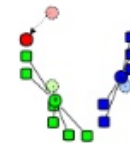
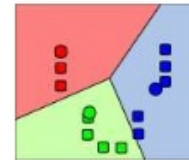where μ_i is the mean of points in S_i.

- ## Standard Algorithm
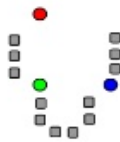
1) k initial "means" (in this case k=3) are randomly selected from the data set.	2) k clusters are created by associating every observation with the nearest mean.	3) The centroid of each of the k clusters becomes the new means.	4) Steps 2 and 3 are repeated until convergence has been reached.

Data Product:

As a data product we have clusters after apply the algorithm, this cluster are very helpful to comparing cities.
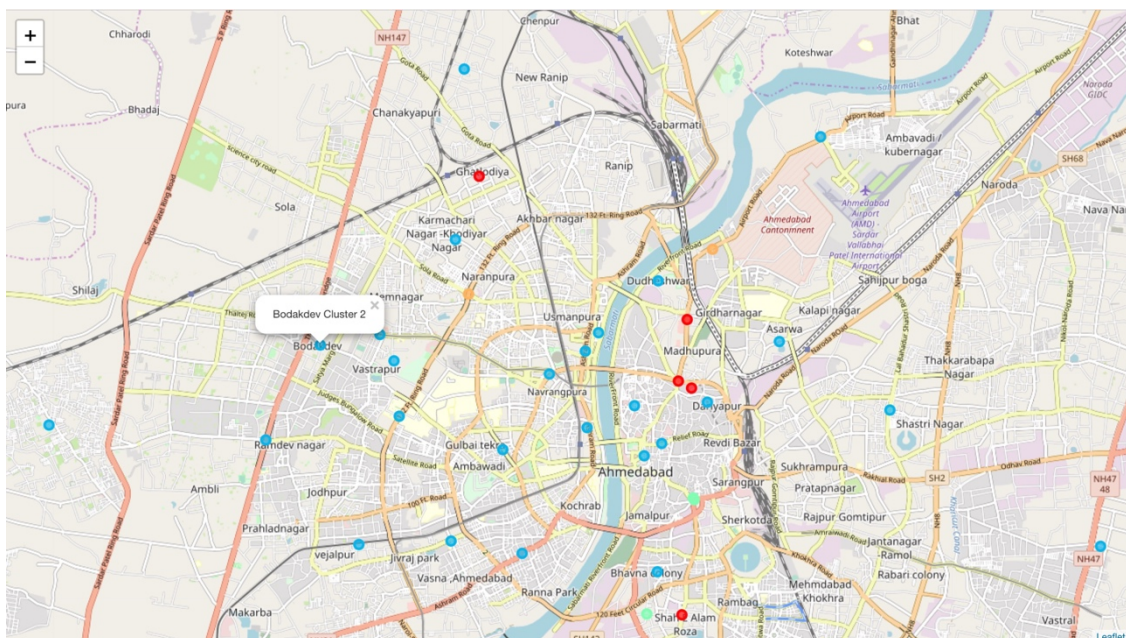
Results

Ahmedabad City Popular Neighborhood's place dataframe and

Clusters:

Out[462]:

|	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
0	Asarwa	23.044980	72.607700	2	Business Service	Tea Room	Women's Store	Department Store	Coffee Shop	Comfort Food Restaurant	Concert Hall	Construction & Landscaping	Convenience Store
1	Ashram Road	23.006986	72.557490	2	Indian Restaurant	Smoke Shop	Department Store	Coffee Shop	Flea Market	Pizza Place	Snack Place	Comfort Food Restaurant	Concert Hall
2	Astodia	23.016830	72.590975	3	Bus Station	Tea Room	Women's Store	Department Store	Coffee Shop	Comfort Food Restaurant	Concert Hall	Construction & Landscaping	Convenience Store
3	Bapunagar	23.032660	72.629494	2	ATM	Arts & Crafts Store	Indian Restaurant	Movie Theater	Food	Food Court	Coffee Shop	Comfort Food Restaurant	Concert Hall
4	Behrampura	23.003589	72.583840	2	Wings Joint	Business Service	Baseball Field	Dessert Shop	Comfort Food Restaurant	Concert Hall	Construction & Landscaping	Convenience Store	Cosmetics Shop
5	Bodakdev	23.044371	72.517921	2	Coffee Shop	Breakfast Spot	Café	Hotel	Tea Room	Indian Restaurant	Bubble Tea Shop	Women's Store	Concert Hall
6	Bopal	23.030050	72.464943	2	Mobile Phone Shop	Park	Fast Food Restaurant	Women's Store	Cricket Ground	Coffee Shop	Comfort Food Restaurant	Concert Hall	Construction & Landscaping
7	CG Road	23.039084	72.562786	2	Indian Restaurant	Pizza Place	Cricket Ground	Food	Fast Food Restaurant	Café	Ice Cream Shop	Clothing Store	Jewelry Store
8	Chandkheda	23.112650	72.583618	2	Shopping Mall	Breakfast Spot	Indian Restaurant	Women's Store	Comfort Food Restaurant	Concert Hall	Construction & Landscaping	Convenience Store	Cosmetics Shop
9	Chandlodia	23.093990	72.546181	2	Snack Place	Women's Store	Coffee Shop	Indian Restaurant	Dessert Shop	Comfort Food Restaurant	Concert Hall	Construction & Landscaping	Convenience Store

Toronto City Popular Neighborhood's place dataframe and

Clusters:

Out[463]:

|	Postcode	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Cor V
0	M4E	East Toronto	The Beaches	43.676357	-79.293031	2	Coffee Shop	Gym / Fitness Center	Pub	Women's Store	Discount Store	Fast Food Restaurant	Farmers Market	Resta
1	M4K	East Toronto	The Danforth West, Riverdale	43.679557	-79.352188	2	Greek Restaurant	Coffee Shop	Ice Cream Shop	Bookstore	Italian Restaurant	Yoga Studio	Cosmetics Shop	Br
2	M4L	East Toronto	The Beaches West, India Bazaar	43.668999	-79.315572	2	Park	Sushi Restaurant	Board Shop	Brewery	Burger Joint	Sandwich Place	Burrito Place	
3	M4M	East Toronto	Studio District	43.659526	-79.340923	1	Café	Coffee Shop	Bakery	Italian Restaurant	American Restaurant	Yoga Studio	Coworking Space	Se Resta
4	M4N	Central Toronto	Lawrence Park	43.728020	-79.388790	2	Bus Line	Park	Swim School	Dim Sum Restaurant	Women's Store	Fast Food Restaurant	Farmers Market	Resta
5	M4P	Central Toronto	Davisville North	43.712751	-79.390197	2	Park	Hotel	Burger Joint	Food & Drink Shop	Dance Studio	Clothing Store	Sandwich Place	Bre
6	M4R	Central Toronto	North Toronto West	43.715383	-79.405678	2	Sporting Goods Shop	Coffee Shop	Clothing Store	Salon / Barbershop	Sandwich Place	Diner	Rental Car Location	Furn
7	M4S	Central Toronto	Davisville	43.704324	-79.388790	2	Pizza Place	Dessert Shop	Sandwich Place	Italian Restaurant	Seafood Restaurant	Café	Sushi Restaurant	C
8	M4T	Central Toronto	Moore Park, Summerhill East	43.689574	-79.383160	2	Playground	Gym	Tennis Court	Park	Farmers Market	Falafel Restaurant	Event Space	Eth Resta
9	M4V	Central Toronto	Deer Park, Forest Hill SE, Rathnelly, South Hi...	43.686412	-79.400049	2	Pub	Coffee Shop	Light Rail Station	Supermarket	Sushi Restaurant	Bagel Shop	Sports Bar	Am Resta
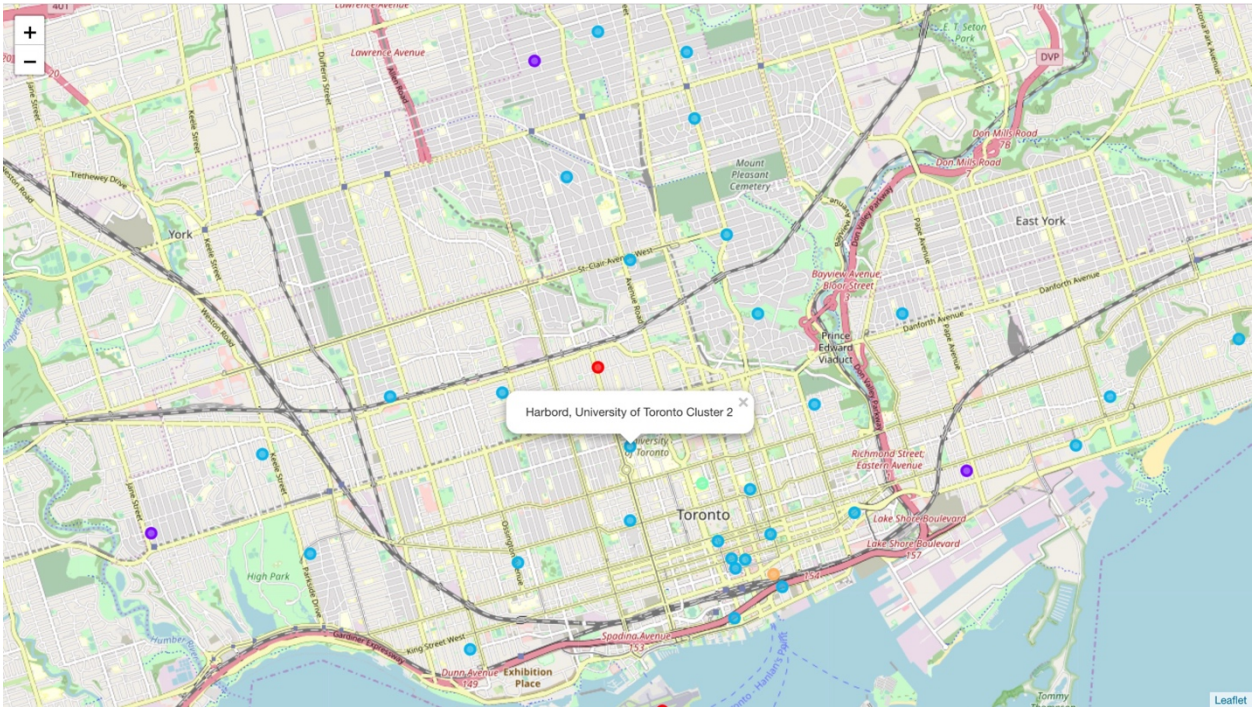
Discussion : For our business problem

1. From cluster and   clustering map of Ahmedabad and Toronto**,** **cluster-3 of Ahmedabad** which also known as **West Ahmedabad** and **cluster-3 of Toronto** are seen like similar in form of Coffey-shop, pizza-Restaurant, Fast-food Restaurant, Café.

2. For Tourism, As per the **Ahmedabad cluster -1** it seems like Tourism place, also there heritage place Delhi Darwaza is situated. It's also have lake in this cluster And the **Ahmedabad cluster-3** seems like Residential Area.

3. For Toronto no data shown for tourism in clusters, all clusters seems like residential area.

4. For Optimal place for Business, **Ahmedabad cluster-3 and Toronto cluster-3** are seems like good option for Business like Coffey-shop, fast-food restaurant, women's store, boutique, Bookstore, pool, Hotel.

5. From the both cluster of Ahmedabad and Toronto we can say that in Vegetarian ,Mexican, Fast-food ,pizza and Coffey shop are very famous.

6. For Optimal location near Bus-station or Railway station,

   a. In Ahmedabad From the Cluster-4 ,we can say that **Dani lambda and Astodia** are best location. There are no data for Bus-station or Railway station in Toronto City.
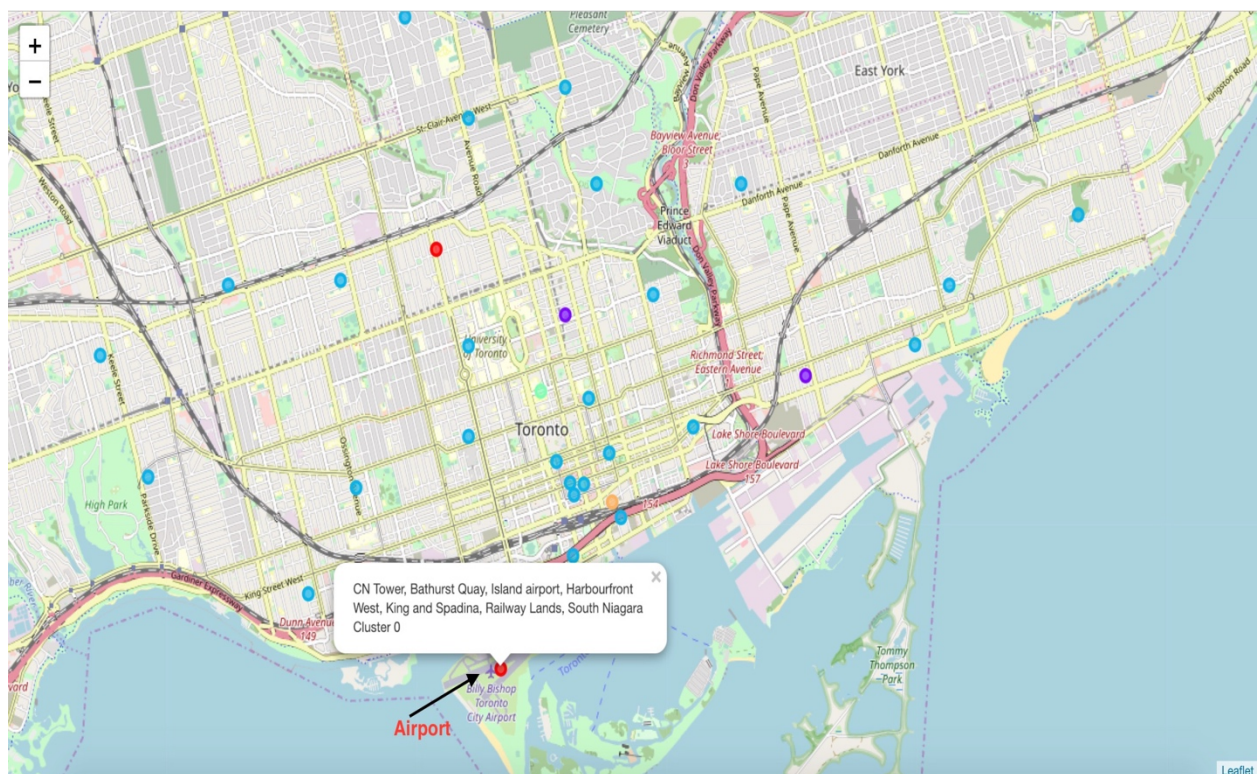
7. For Airport Facility,

   a. In Ahmedabad city, from the **Cluster-2 Ahmedabad** we can say that **Sardar nagar** is near from the **Airport** and Best
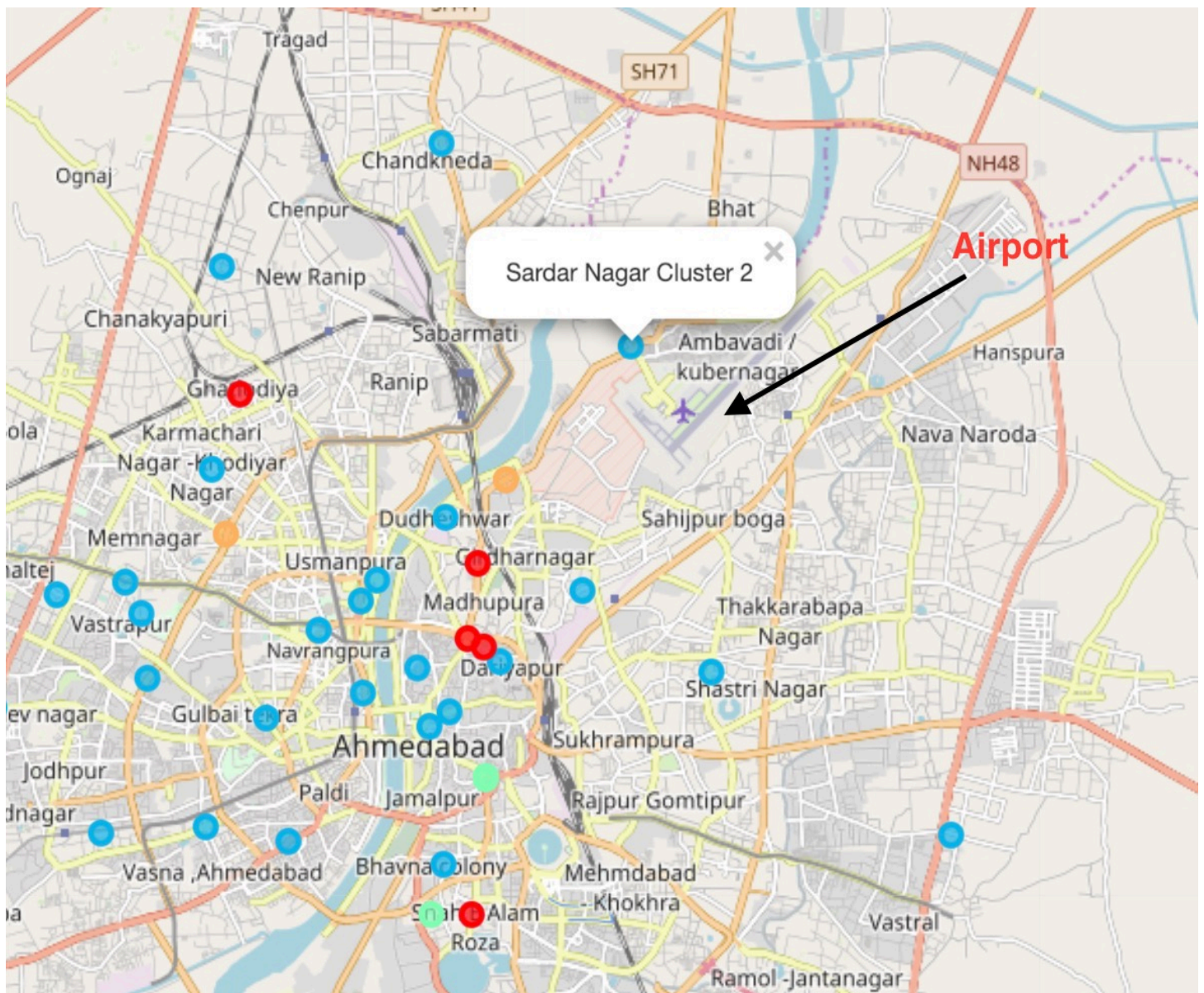
Option for Hotel Business, because very few hotel around

them.

b. In Toronto City, From the Toronto-City-Cluster 1 we can say

that **CN Tower, Bathurst Quay** are best for airport facility,

and also airport longue is there.

Conclusion

From the about discussion about our project we can say that,

- West Ahmedabad is little bit similar like Toronto.

- Ahmedabad have good heritage places and also have Residential Area, and all Toronto area are look like Residential Area.

- Good business option in Ahmedabad and Toronto are Coffey-shop, fast-food restaurant, women's store, boutique, Bookstore, pool, Café.

- Vegetarian, Mexican, Fast-food, Pizza and Chinese food are famous in Ahmedabad and Canada.

- For Bus and Railway Station Dani lambda and Astodia are optimal location in Ahmedabad.

- For Airport Facility, CN Tower, Bathurst Quay in **Toronto** and Sardar Nagar in **Ahmedabad** are optimal place.

- **We can say that, Both Toronto and Ahmedabad City are optimal place for Business**.

THNAKS FOR READING

BY: MOHIT RAJPUT