

# Capstone Project – Battle of Neighborhoods – Ahmedabad

## Introduction

Ahmedabad is metro city of Gujarat, India. It has population of more than 6.35 million and area of 464.16km<sup>2</sup>. As it's the 7<sup>th</sup> largest city of India by population, it provides immense opportunities for all sort of businesses. But Ahmedabad is divided into 29 postal zones and they all are different in terms of opportunities.

Here, I am trying to identify similar neighborhoods in Ahmedabad city using analytical approach and machine learning techniques. This solution can help executives to guide their business or help individuals to move in city. So, this project answers the questions like “Which neighborhoods are similar in terms of venues?” “Where can I open new franchise, which has similarity of my current area of business?” or “Where can I move in city which is very similar to the place I live?”

## Data

First, Geographical data of Ahmedabad is needed. All the pincodes and areas in it are required. This data can be downloaded from India census 2011 website. This file consists all the pincode data of India. We need to filter out the pincodes of Ahmedabad City. According to Wikipedia, Ahmedabad City's pincode starts with 380XXX. So, I filtered out all the pincodes from 380000 to 380999.

Many areas are divided by multiple pincodes. We need unique pincodes. In order to achieve uniqueness of pincodes, I combined area names with same pincodes. It gives me 46 unique pincodes. But out of 46 pincodes, some are outside main city. Those are in outskirts or far away from main city. We need to focus on main city only. So, I removed pincodes which doesn't have Ahmedabad City in Borough Column.

Main City has 29 pincodes. Now we need coordinates of all the pincodes so that we can use them to get foursquare data using API. I used Geocoder package to access Google Geolocation API on order to get coordinates.

Using foursquare API, we can get venues in each pincode area based on search radius of 1100m.

## Methodology

Using Foursquare API, we can get the venues in each pincode area. We have found 360 venues in total.

```
print(Ahmedabad_venues.shape)
Ahmedabad_venues.head()
```

(360, 7)

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Ahmedabad G.P.O., Lal Darwaja S.O., Raipur S.O ...	23.026426	72.584341	Manek Chowk Khau Gali	23.023505	72.588539	Snack Place
1	Ahmedabad G.P.O., Lal Darwaja S.O., Raipur S.O ...	23.026426	72.584341	Manek Chowk	23.023626	72.588553	Fast Food Restaurant
2	Ahmedabad G.P.O., Lal Darwaja S.O., Raipur S.O ...	23.026426	72.584341	Lucky Tea	23.027829	72.581394	Tea Room
3	Ahmedabad G.P.O., Lal Darwaja S.O., Raipur S.O ...	23.026426	72.584341	Agashiye	23.027104	72.581614	Indian Restaurant
4	Ahmedabad G.P.O., Lal Darwaja S.O., Raipur S.O ...	23.026426	72.584341	House of MG	23.027186	72.581621	Hotel

The table below shows the count of venues found in each area with given condition.

Ahmedabad_venues.groupby('Neighborhood').count()						
	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Neighborhood						
Ahmedabad G.P.O., Lal Darwaja S.O, Raipur S.O (Ahmedabad), Gheekanta Road S.O, Manek Chowk S.O, Dariapur S.O (Ahmedabad), Gandhi Road S.O (Ahmedabad), Khadia S.O, Raikhad S.O, District Court S.O (Ahmedabad), Khanpur S.O (Ahmedabad), Jamalpur S.O (Ahmedabad), Municipal Corporation S.O	13	13	13	13	13	13
Amraiwadi S.O, Ctm Char Rasta S.O	6	6	6	6	6	6
Asarwa Ext South S.O, I E Bapunagar S.O, Bapunagar S.O	5	5	5	5	5	5
Bodakdev S.O, Thaltej Road S.O	12	12	12	12	12	12
Calico Mills S.O, M D Marg S.O, Behrampura S.O, Gita Mandir Road S.O	2	2	2	2	2	2
D Cabin S.O, Railway Colony S.O (Ahmedabad)	6	6	6	6	6	6
Ellisbridge S.O	39	39	39	39	39	39
Gandhi Ashram S.O (Ahmedabad)	11	11	11	11	11	11
Ghatlodia S.O	6	6	6	6	6	6
Ghodasar S.O (Ahmedabad)	5	5	5	5	5	5
Jivraj Park S.O	1	1	1	1	1	1
Jodhpur Char Rasta S.O, Azad Society S.O, I I M S.O, Polytechnic S.O (Ahmedabad), Manekbag S.O, S A C S.O	17	17	17	17	17	17
Juhapura S.O	4	4	4	4	4	4
Kabir Chowk S.O, Sabarmati S.O, O N G C S.O	4	4	4	4	4	4
Maninagar S.O, Daxini Society S.O, Khokhara Mehmabad S.O, L G Hospital S.O, Vasisthnagar S.O, Jawahar Chowk S.O, S A Mills S.O	12	12	12	12	12	12
Meghaningar S.O, Public Office S.O (Ahmedabad), Asarwa Chakla S.O, Civil Hospital S.O	5	5	5	5	5	5
Memnagar S.O	25	25	25	25	25	25
N C Market S.O, Revdibazar H.O, Railwaypura S.O	10	10	10	10	10	10
Navjivan S.O	26	26	26	26	26	26
Navrangpura H.O, Gujarat University S.O	55	55	55	55	55	55
Nikol S.O	40	40	40	40	40	40
Rajpur Gomtipur S.O, Gomtipur S.O	1	1	1	1	1	1
Saraspur S.O	6	6	6	6	6	6
Sarkhej Road S.O, Anandnagar S.O (Ahmedabad), Shardanagar S.O, Narayannagar S.O, Paldi S.O (Ahmedabad)	6	6	6	6	6	6
Shah Alam Roza S.O, Bhairavnath Road S.O	5	5	5	5	5	5
Sola H B C S.O	10	10	10	10	10	10
Sub Foreign S.O, Dudheshwar Tavdipura S.O, Girdharnagar S.O, Cantonment S.O, Delhi Gate S.O (Ahmedabad), Shahpur S.O (Ahmedabad), Madhupura Market S.O, Shahibag S.O	8	8	8	8	8	8
Sukhrampura S.O, Rakhial Udyog Vistar S.O, Rakhial S.O	3	3	3	3	3	3
Vadaj S.O, Naranpura Vistar S.O, Stadium Marg S.O, Shastrinagar S.O (Ahmedabad), Nava Vadaj S.O	17	17	17	17	17	17

As one can notice, many areas have very less venues. We can eliminate them later as they don't have enough data to analyze. There are total 96 unique categories. One hot encoding can be used to get frequency of each venue with respective to pincode area.

	Neighborhood_names	ATM	Accessories Store	Arcade	Arts & Crafts Store	Asian Restaurant	BBQ Joint	Bakery	Basketball Court	Bistro	...	Supe
0	Ahmedabad G.P.O., Lal Darwaja S.O, Raipur S.O ...	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000	0.0	0.0	...	0.0
1	Amraiwadi S.O, Ctm Char Rasta S.O	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000	0.0	0.0	...	0.0
2	Asarwa Ext South S.O, I E Bapunagar S.O, Bapun...	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000	0.0	0.0	...	0.0
3	Bodakdev S.O, Thaltej Road S.O	0.0	0.0	0.0	0.0	0.0	0.083333	0.083333	0.0	0.0	...	0.0
4	Calico Mills S.O, M D Marg S.O, Behrampura S.O...	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000	0.0	0.0	...	0.0

5 rows × 96 columns

Based on this result we can find the top venues in each area. Below is the report for top 5 venues in each area.

----Ahmedabad G.P.O., Lal Darwaja S.O, Raipur S.O (Ahmedabad), Gheekanta Road S.O, Manek Chowk S.O, Dariapur S.O (Ahmedabad), Gandhi Road S.O (Ahmedabad), Khadia S.O, Raikhad S.O, District Court S.O (Ahmedabad), Khanpur S.O (Ahmedabad), Jamalpur S.O (Ahmedabad), Municipal Corporation S.O----

	venue	freq
0	Hotel	0.31
1	Historic Site	0.15
2	Indian Restaurant	0.15
3	Vegetarian / Vegan Restaurant	0.08
4	Tea Room	0.08

----Amraiwadi S.O, Ctm Char Rasta S.O----

	venue	freq
0	Pool Hall	0.17
1	IT Services	0.17
2	Moving Target	0.17
3	Indian Restaurant	0.17
4	Brewery	0.17

----Asarwa Ext South S.O, I E Bapunagar S.O, Bapunagar S.O----

	venue	freq
0	Clothing Store	0.2
1	Indian Restaurant	0.2
2	Gourmet Shop	0.2
3	Movie Theater	0.2
4	Bridal Shop	0.2

----Bodakdev S.O, Thaltej Road S.O----

	venue	freq
0	Hotel	0.17
1	Indian Restaurant	0.17
2	Tennis Court	0.08
3	Fast Food Restaurant	0.08
4	Event Space	0.08

----Calico Mills S.O, M D Marg S.O, Behrampura S.O, Gita Mandir Road S.O--  
--

	venue	freq
0	Cafeteria	0.5
1	Boat or Ferry	0.5
2	ATM	0.0
3	Mexican Restaurant	0.0
4	Pharmacy	0.0

----D Cabin S.O, Railway Colony S.O (Ahmedabad)----

	venue	freq
0	Ice Cream Shop	0.50
1	Farmers Market	0.17
2	Bus Station	0.17
3	Pizza Place	0.17
4	ATM	0.00

----Ellisbridge S.O----

	venue	freq
0	Fast Food Restaurant	0.18
1	Café	0.13
2	Indian Restaurant	0.08
3	Shopping Mall	0.05
4	Mexican Restaurant	0.05

----Gandhi Ashram S.O (Ahmedabad)----

	venue	freq
0	Pizza Place	0.09
1	Snack Place	0.09
2	Hotel Bar	0.09
3	Indian Restaurant	0.09
4	Indie Movie Theater	0.09

----Ghatlodia S.O----

	venue	freq
0	Indian Restaurant	0.33
1	Pizza Place	0.17
2	Arts & Crafts Store	0.17
3	Ice Cream Shop	0.17
4	Café	0.17

----Ghodasar S.O (Ahmedabad)----

	venue	freq
0	ATM	0.4
1	Indian Restaurant	0.2
2	Motorcycle Shop	0.2
3	Pizza Place	0.2
4	Mexican Restaurant	0.0

----Jivraj Park S.O----

	venue	freq
0	Fried Chicken Joint	1.0
1	ATM	0.0
2	Men's Store	0.0
3	Park	0.0
4	Noodle House	0.0

----Jodhpur Char Rasta S.O, Azad Society S.O, I I M S.O, Polytechnic S.O (Ahmedabad), Manekbag S.O, S A C S.O----

	venue	freq
0	Café	0.18
1	Fast Food Restaurant	0.12
2	Clothing Store	0.12
3	Mexican Restaurant	0.06
4	Pizza Place	0.06

----Juhapura S.O----

	venue	freq
0	IT Services	0.25
1	Fried Chicken Joint	0.25
2	Food & Drink Shop	0.25
3	Middle Eastern Restaurant	0.25
4	ATM	0.00

----Kabir Chowk S.O, Sabarmati S.O, O N G C S.O----

	venue	freq
0	Fast Food Restaurant	0.25
1	Multiplex	0.25
2	Sandwich Place	0.25
3	Cricket Ground	0.25
4	Ice Cream Shop	0.00

----Maninagar S.O, Daxini Society S.O, Khokhara Mehmadabad S.O, L G Hospital S.O, Vasisthnagar S.O, Jawahar Chowk S.O, S A Mills S.O----

	venue	freq
0	Fast Food Restaurant	0.33
1	Pizza Place	0.08
2	Accessories Store	0.08

3	Ice Cream Shop	0.08
4	Indian Restaurant	0.08

----Meghaningar S.O, Public Office S.O (Ahmedabad), Asarwa Chakla S.O, Civil Hospital S.O----

	venue	freq
0	Indian Restaurant	0.4
1	Stadium	0.2
2	Pharmacy	0.2
3	History Museum	0.2
4	IT Services	0.0

----Memnagar S.O----

	venue	freq
0	Hotel	0.12
1	Fast Food Restaurant	0.12
2	Pizza Place	0.12
3	Café	0.08
4	Indian Restaurant	0.08

----N C Market S.O, Revdibazar H.O, Railwaypura S.O----

	venue	freq
0	Shopping Mall	0.2
1	Zoo	0.1
2	Sports Club	0.1
3	Fast Food Restaurant	0.1
4	Multiplex	0.1

----Navjivan S.O----

	venue	freq
0	Café	0.15
1	Indian Restaurant	0.12
2	Ice Cream Shop	0.08
3	Asian Restaurant	0.08
4	Pizza Place	0.08

----Navrangpura H.O, Gujarat University S.O----

	venue	freq
0	Indian Restaurant	0.18
1	Hotel	0.13
2	Fast Food Restaurant	0.07
3	Dessert Shop	0.05
4	Coffee Shop	0.05

----Nikol S.O----

	venue	freq
0	Chinese Restaurant	0.18
1	Food Court	0.08

2	Café	0.08
3	Hostel	0.05
4	BBQ Joint	0.05

----Rajpur Gomtipur S.O, Gomtipur S.O----

	venue	freq
0	Indian Restaurant	1.0
1	ATM	0.0
2	Men's Store	0.0
3	Park	0.0
4	Noodle House	0.0

----Saraspur S.O----

	venue	freq
0	Food Court	0.17
1	Asian Restaurant	0.17
2	Indian Restaurant	0.17
3	Bakery	0.17
4	Clothing Store	0.17

----Sarkhej Road S.O, Anandnagar S.O (Ahmedabad), Shardanagar S.O, Narayan nagar S.O, Paldi S.O (Ahmedabad)----

	venue	freq
0	Hotel	0.17
1	Coffee Shop	0.17
2	Ice Cream Shop	0.17
3	Indian Restaurant	0.17
4	Snack Place	0.17

----Shah Alam Roza S.O, Bhairavnath Road S.O----

	venue	freq
0	Food & Drink Shop	0.2
1	Lake	0.2
2	Bus Stop	0.2
3	Food Court	0.2
4	Bus Station	0.2

----Sola H B C S.O----

	venue	freq
0	Indian Restaurant	0.3
1	Pizza Place	0.2
2	Department Store	0.1
3	Gym	0.1
4	Fast Food Restaurant	0.1

----Sub Foreign S.O, Dudheshwar Tavdipura S.O, Girdharnagar S.O, Cantonment S.O, Delhi Gate S.O (Ahmedabad), Shahpur S.O (Ahmedabad), Madhupura Market S.O, Shahibag S.O----

```

          venue  freq
0   History Museum  0.12
1   Ice Cream Shop  0.12
2  Indian Restaurant  0.12
3           Bakery  0.12
4           Museum  0.12

```

----Sukhrampura S.O, Rakhial Udyog Vistar S.O, Rakhial S.O----

```

          venue  freq
0  Fast Food Restaurant  0.33
1      Gourmet Shop  0.33
2   Indian Restaurant  0.33
3      Men's Store  0.00
4           Park  0.00

```

----Vadaj S.O, Naranpura Vistar S.O, Stadium Marg S.O, Shastrinagar S.O (Ahmedabad), Nava Vadaj S.O----

```

          venue  freq
0   Indian Restaurant  0.18
1      Pizza Place  0.12
2   Sandwich Place  0.12
3  Fast Food Restaurant  0.06
4      Food Court  0.06

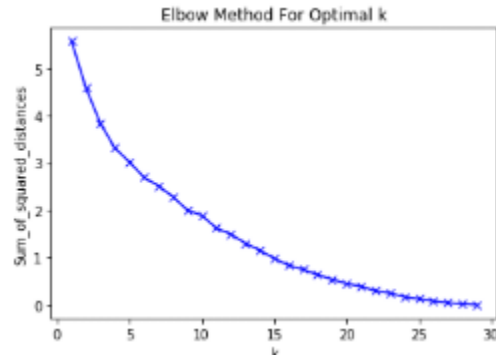
```

Based on this information, we can rank the top 10 venues in each neighborhood. Result is something like this.

	Neighborhood_names	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	Ahmedabad G.P.O., Lal Darwaja S.O, Raipur S.O ...	Hotel	Historic Site	Indian Restaurant	Tea Room	Fast Food Restaurant	Snack Place	Park	Vegetarian / Vegan Restaurant	Fala Rest
1	Amraiwadi S.O, Ctm Char Rasta S.O	Brewery	Health Food Store	Moving Target	Pool Hall	IT Services	Indian Restaurant	Event Space	Cricket Ground	Dep: Ston
2	Asanwa Ext South S.O, I E Bapunagar S.O, Bapun...	Bridal Shop	Gourmet Shop	Clothing Store	Movie Theater	Indian Restaurant	Zoo	Cricket Ground	Department Store	Des: Sho
3	Bodakdev S.O, Thaltej Road S.O	Hotel	Indian Restaurant	Fast Food Restaurant	Event Space	Clothing Store	BBQ Joint	Tennis Court	Bakery	Café
4	Calico Mills S.O, M D Marg S.O, Behrampura S.O...	Cafeteria	Boat or Ferry	Zoo	Falafel Restaurant	Cricket Ground	Department Store	Dessert Shop	Dim Sum Restaurant	Dine

We can use this data to identify the similar areas. That's when unsupervised machine learning comes in picture. I have decided to go with kmeans clustering as I want to know the similarity and common trend to define similar areas. To decide the number of clusters, I am going with elbow method. As I have 29 observations in the data, I can select clusters between 1 to 29. But to find the optimum number of clusters I have plotted total within-cluster sum of square vs number of clusters.





As you can see, error goes down as number of clusters increases. But after k=12 the slope decreases as number of clusters increases. It is called elbow method. After kmeans clustering with 12 clusters we can use the cluster labels to the area.

## Results

Out of 12 clusters only 3 clusters have more than one pincode. They are mentioned below:

Cluster 2:

Cluster 2

```
ahm_merged.loc[ahm_merged['Cluster Labels'] == 1, ahm_merged.columns[[0,2] + list(range(5, ahm_merged.shape[1]))]]
```

	Postcode	Neighborhood_names	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
1	380002	N C Market S.O, Revdibazar H.O, Railway pura S.O	1	Shopping Mall	Zoo	Fast Food Restaurant	Sandwich Place	Bus Station	Burger Joint	Multiplex
10	380015	Jodhpur Char Rasta S.O, Azad Society S.O, II ...	1	Café	Clothing Store	Fast Food Restaurant	Snack Place	Coffee Shop	Mexican Restaurant	Pizza Place
19	380027	Gandhi Ashram S.O (Ahmedabad)	1	Furniture / Home Store	Pizza Place	Sculpture Garden	Indie Movie Theater	Indian Restaurant	Hotel Bar	Fast Food Restaurant
21	380049	Nikol S.O	1	Chinese Restaurant	Food Court	Café	Noodle House	Vegetarian / Vegan Restaurant	BBQ Joint	Dim Sum Restaurant

Cluster 4

Cluster 4

```
ahm_merged.loc[ahm_merged['Cluster Labels'] == 3, ahm_merged.columns[[0,2] + list(range(5, ahm_merged.shape[1]))]]
```

	Postcode	Neighborhood_names	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
0	380001	Ahmedabad G.P.O., Lal Darwaja S.O, Raipur S.O ...	3	Hotel	Historic Site	Indian Restaurant	Tea Room	Fast Food Restaurant	Snack Place	Park
2	380004	Sub Foreign S.O, Dudheshwar Tavdipura S.O, Gir...	3	Museum	Garden	Furniture / Home Store	History Museum	Pizza Place	Bakery	Ice Cream Shop
4	380006	Ellisbridge S.O	3	Fast Food Restaurant	Café	Indian Restaurant	Shopping Mall	Tea Room	Park	Pizza Place
6	380008	Maninagar S.O, Daxini Society S.O, Khokhara Me...	3	Fast Food Restaurant	Indian Restaurant	Tennis Court	Pizza Place	Electronics Store	Ice Cream Shop	Clothing Store
7	380009	Navrangpura H.O, Gujarat University S.O	3	Indian Restaurant	Hotel	Fast Food Restaurant	Café	Snack Place	Dessert Shop	Coffee Shop
9	380014	Navjivan S.O	3	Café	Indian Restaurant	Asian Restaurant	Pizza Place	Ice Cream Shop	Cricket Ground	Dessert Shop
12	380018	Saraspur S.O	3	Clothing Store	Asian Restaurant	Bakery	Food Court	Indian Restaurant	Men's Store	Zoo
17	380024	Asarwa Ext South S.O, I E Bapunagar S.O, Bapun...	3	Bridal Shop	Gourmet Shop	Clothing Store	Movie Theater	Indian Restaurant	Zoo	Cricket Ground
18	380026	Amraiwadi S.O, Ctm Char Rasta S.O	3	Brewery	Health Food Store	Moving Target	Pool Hall	IT Services	Indian Restaurant	Event Space
24	380052	Memnagar S.O	3	Pizza Place	Fast Food Restaurant	Hotel	Indian Restaurant	Chinese Restaurant	Café	Shopping Mall
25	380054	Bodakdev S.O, Thaltej Road S.O	3	Hotel	Indian Restaurant	Fast Food Restaurant	Event Space	Clothing Store	BBQ Joint	Tennis Court

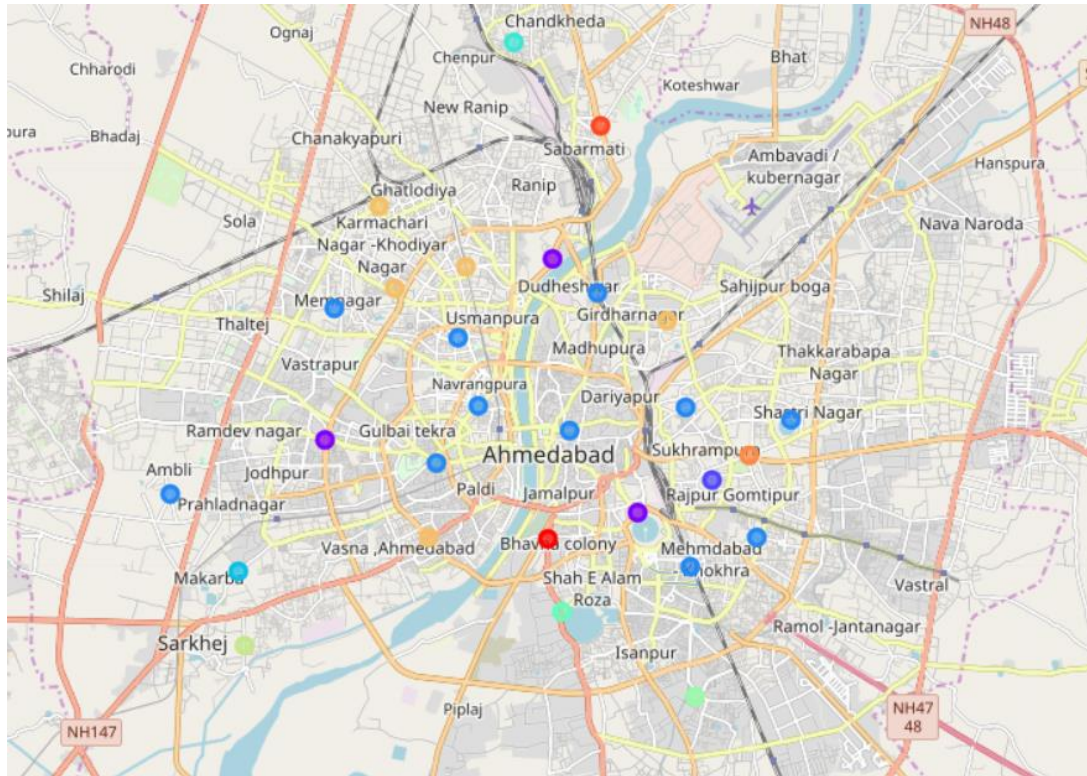
## Cluster 10:

Cluster 10

```
ahm_merged.loc[ahm_merged['Cluster Labels'] == 9, ahm_merged.columns[[0,2] + list(range(5, ahm_merged.shape[1]))]]
```

	Postcode	Neighborhood_names	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
5	380007	Sarkhej Road S.O, Anandnagar S.O (Ahmedabad), ...	9	Pizza Place	Hotel	Ice Cream Shop	Indian Restaurant	Coffee Shop	Snack Place	Zoo
8	380013	Vadaj S.O, Naranpura Vistar S.O, Stadium Marg ...	9	Indian Restaurant	Sandwich Place	Pizza Place	Gym	Bus Station	Ice Cream Shop	Dessert Shop
11	380016	Meghaningar S.O, Public Office S.O (Ahmedabad)...	9	Indian Restaurant	Pharmacy	History Museum	Stadium	Event Space	Convenience Store	Cosmeti Shop
27	380061	Ghatlodia S.O	9	Indian Restaurant	Pizza Place	Arts & Crafts Store	Café	Ice Cream Shop	Zoo	Event Space
28	380063	Sola H B C S.O	9	Indian Restaurant	Pizza Place	Gym	Vegetarian / Vegan Restaurant	Hotel	Department Store	Fast Food Restaurant

It can be visualized in map of Ahmedabad.



## **Discussion**

As one can notice, the number of venues we get from foursquare doesn't reflect the reality, as Ahmedabad has far more venues. But, since Foursquare is not popular in India, not many venues are registered on foursquare.

## **Conclusion:**

The results suggest that this approach would yield insights on common trends between areas. This analysis can be further improved by using more popular source like Google Maps and adding Demographics from India Census 2011