

# Data

The data acquired from Wikipedia pages and restructure to csv file for easier manipulation and reading. Both files uploaded to my GitHub for references. Link to the files are:

- [Mumbai district.csv](#)
  - [https://github.com/skakkara-in-ibm-com/Coursera\\_Capstone\\_Assignment1/blob/master/Mumbai\\_district.csv](https://github.com/skakkara-in-ibm-com/Coursera_Capstone_Assignment1/blob/master/Mumbai_district.csv)
- [Pune district.csv](#)
  - [https://github.com/skakkara-in-ibm-com/Coursera\\_Capstone\\_Assignment1/blob/master/Pune\\_district.csv](https://github.com/skakkara-in-ibm-com/Coursera_Capstone_Assignment1/blob/master/Pune_district.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.

To start, let's get and look at the data. I've already downloaded it, so let's read it (from local drive) and load it to dataframe: Using geocoder, we able to get Latitude and longitude for each area.

```
[5]: #now, using Geocoder and Google API, we get the Latitude and Longitude of each area
import geocoder
SSK_API_KEY='AIzaSyCGMwm7FMlp7zamQX-LJOjckjvFb5MYxk'
#function to get Latitude and Longitude
def get_latlng(postal_code):
    lat_lng_coords = None
    while(lat_lng_coords is None):
        g = geocoder.google('{}, India'.format(postal_code), key=SSK_API_KEY)
        #print(g)
        lat_lng_coords = g.latlng
    return lat_lng_coords
```

```
[6]: #put new column of Latitude and Longitude into dataframe
postal_codes_mum = df_mum['Taluka']+df_mum['Area']
coords = [ get_latlng(postal_code) for postal_code in postal_codes_mum.tolist() ]
df_mum_coords = pd.DataFrame(coords, columns=['Latitude', 'Longitude'])
df_mum['Latitude'] = df_mum_coords['Latitude']
df_mum['Longitude'] = df_mum_coords['Longitude']
df_mum.head(10)
```

	Postcode	District	Area	Taluka	PostalStatus	Latitude	Longitude
0	400001	Mumbai	Bazargate	Mumbai	Sub Office	18.936518	72.835823
1	400001	Mumbai	M P T	Mumbai	Sub Office	18.948288	72.844163
2	400001	Mumbai	Stock Exchange	Mumbai	Sub Office	18.930012	72.833639
3	400001	Mumbai	Tajmahal	Mumbai	Sub Office	18.921729	72.833031
4	400001	Mumbai	Town Hall Mumbai	Mumbai	Sub Office	19.115063	72.850029
5	400001	Mumbai	Mumbai G P O	Mumbai	Head Office	18.938956	72.837057
6	400002	Mumbai	Kalbadavi	Mumbai	Head Office	18.948738	72.828883
7	400002	Mumbai	Ramwadi	Mumbai	Sub Office	18.950313	72.827458
8	400002	Mumbai	S C Court	Mumbai	Sub Office	18.931120	72.830150
9	400002	Mumbai	Thakurdwar	Mumbai	Sub Office	18.949847	72.822745

```
[7]: #new column for Pune dataframe
postal_codes_pun = df_pun['Taluka']+df_pun['Area']
coords = [ get_latlng(postal_code) for postal_code in postal_codes_pun.tolist() ]
df_pun_coords = pd.DataFrame(coords, columns=['Latitude', 'Longitude'])
df_pun['Latitude'] = df_pun_coords['Latitude']
df_pun['Longitude'] = df_pun_coords['Longitude']
df_pun.head(10)
```

	Postcode	District	Area	Taluka	PostalStatus	Latitude	Longitude
0	410301	Pune	Khandala	Maval	Sub Office	18.755211	73.444539
1	410302	Pune	R P T S Khandala	Maval	Sub Office	18.769203	73.376764
2	410401	Pune	Ambavane	Maval	Branch Office	20.593684	78.962880
3	410401	Pune	Kusgaon Bk	Maval	Branch Office	20.593684	78.962880
4	410401	Pune	Lonavala Bazar	Maval	Sub Office	20.593684	78.962880
5	410401	Pune	Lonavala	Maval	Sub Office	18.755211	73.444539
6	410402	Pune	Ins Shivaji Lonavale	Maval	Sub Office	18.746499	73.399671
7	410402	Pune	Kurwande	Maval	Branch Office	20.593684	78.962880
8	410403	Pune	Kaivalyadham	Maval	Sub Office	18.760702	73.415582
9	410405	Pune	Kamshet	Maval	Sub Office	18.755211	73.444539

# Methodology

In this project, I will use the basic methodology as taught in Week 3 lab.

- Above, we have done convert addresses into their equivalent latitude and longitude values.
- Then we will use the Foursquare API to explore neighbourhoods in both cities, Mumbai and Pune
- After that, explore function to get the most common venue categories in each neighbourhood,
- and then use this feature to group the neighbourhoods into clusters

K-means clustering algorithm will be using to complete this task. And the Folium library to visualize the neighbourhoods in Mumbai and Pune and their emerging clusters.

Based on dataframe analysis above, we found out that **Mumbai Taluka** in **Mumbai** and **Pune City Taluka** in **Pune** are both have the highest number of area within it those districts.

Kurla Netajinagar Nitie Chakala Midc Seepz Vidyanagari Airport Mumbai International Airport Sahar P T Colony Sahargaan (2101, 7)							
[16]:	Area	Area Latitude	Area Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Bazargate	18.936518	72.835823	Royal China	18.938715	72.832933	Chinese Restaurant
1	Bazargate	18.936518	72.835823	Town House Cafe	18.938550	72.833464	Bar
2	Bazargate	18.936518	72.835823	Dakshin Bar And Kitchen	18.936489	72.837490	Seafood Restaurant
3	Bazargate	18.936518	72.835823	Ideal Corner	18.934961	72.834050	Indian Restaurant
4	Bazargate	18.936518	72.835823	Sher-E-Punjab	18.937944	72.837853	Indian Restaurant
N I B M Anandnagar Pune Karvenagar Navsahyadri Wakad Warje Mohamadwadi Rupeenagar Talwade (911, 7)							
[17]:	Area	Area Latitude	Area Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	C D A O	18.515806	73.927164	Cafe Coffee Day	18.512372	73.928695	Coffee Shop
1	C D A O	18.515806	73.927164	Deccan Harvest	18.514183	73.927503	Indian Restaurant
2	C D A O	18.515806	73.927164	Barbeque Nation	18.517125	73.923614	BBQ Joint
3	Dr B A Chowk	18.523306	73.869719	Coconut Grove	18.525681	73.866989	Seafood Restaurant
4	Dr B A Chowk	18.523306	73.869719	Mahesh Lunch Home	18.520633	73.872360	Seafood Restaurant

# Analyse Mumbai

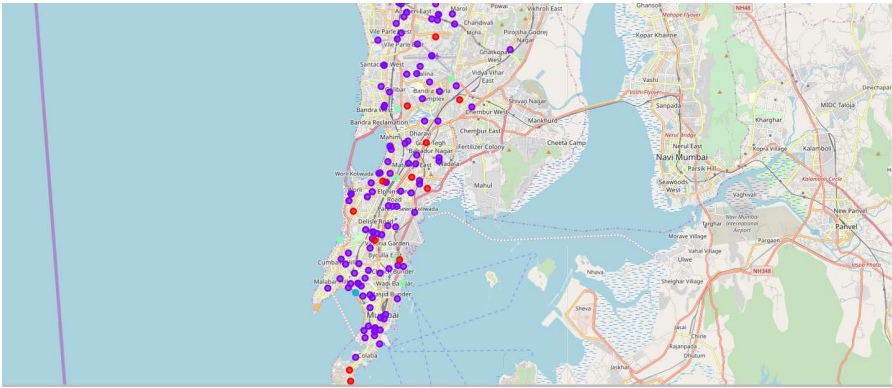
Analysing Mumbai data to get most common venue for each area.

	Area	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
0	A I Staff Colony	Indian Restaurant	Pub	Fast Food Restaurant	Bus Station	Restaurant	Multiplex	Yoga Studio	Fish Market
1	Agripada	Bakery	Indian Restaurant	Gym	Coffee Shop	Restaurant	Yoga Studio	Event Space	Flea Market
2	Airport Mumbai	Airport	Duty-free Shop	Event Space	Food	Flower Shop	Flea Market	Fish Market	Field
3	Ambewadi Mumbai	Juice Bar	Fast Food Restaurant	Breakfast Spot	Vegetarian / Vegan Restaurant	Gym / Fitness Center	Bus Station	Snack Place	Coffee Shop
4	Andheri	Hotel	Indian Restaurant	Seafood Restaurant	Vegetarian / Vegan Restaurant	Hotel Bar	Pizza Place	Department Store	Yoga Studio

# K-mean Cluster Mumbai

Using K-mean to clustering data area with most common venue

Postcode	District	Area	Taluka	PostalStatus	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	
0	400001	Mumbai	Bazargate	Mumbai	Sub Office	18.936518	72.835823	1	Indian Restaurant	Café	Seafood Restaurant	Chinese Restaurant	Irani Cafe	Hotel	Bakery	Park
1	400001	Mumbai	M P T	Mumbai	Sub Office	18.948288	72.844163	1	Harbor / Marina	Hotel	Convenience Store	Cupcake Shop	Cricket Ground	Food	Flower Shop	Flea Market
2	400001	Mumbai	Stock Exchange	Mumbai	Sub Office	18.930012	72.833639	1	Indian Restaurant	Café	Seafood Restaurant	Bookstore	Dessert Shop	Chinese Restaurant	Coffee Shop	Clothing Store
3	400001	Mumbai	Tajmahal	Mumbai	Sub Office	18.921129	72.833031	1	Café	Bar	Hotel	Coffee Shop	Indian Restaurant	Cocktail Bar	Pub	Mughlai Restaurant
4	400001	Mumbai	Town Hall Mumbai	Mumbai	Sub Office	19.115063	72.850029	1	Indian Restaurant	Hotel	Food Court	Burger Joint	Pizza Place	Platform	Café	Chinese Restaurant



# Analyse Pune

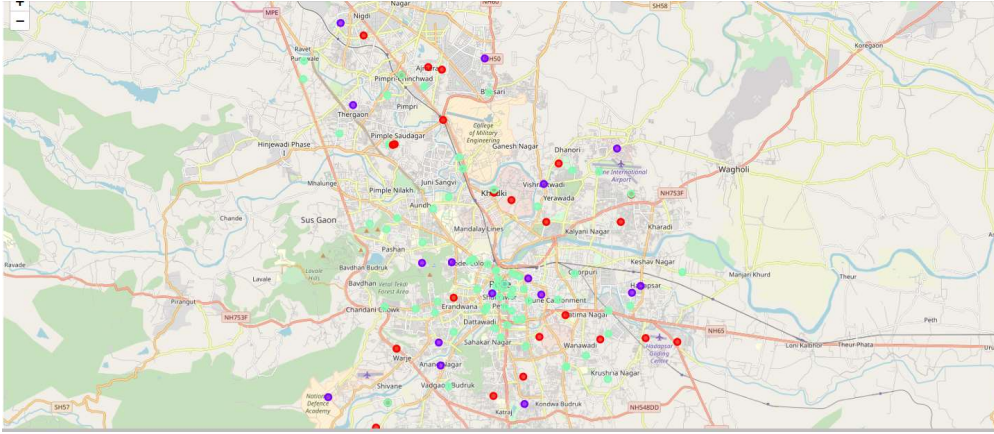
Analysing Pune data to get most common venue for each area.

	Area	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
0	9 Drd	Historic Site	Department Store	Bakery	Women's Store	Falafel Restaurant	Food & Drink Shop	Food	Flea Market
1	A R Shala	Historic Site	Department Store	Bakery	Women's Store	Falafel Restaurant	Food & Drink Shop	Food	Flea Market
2	Afmc	Racetrack	Cupcake Shop	Track Stadium	Women's Store	Dumpling Restaurant	Food	Flea Market	Fast Food Restaurant
3	Airport Pune	Airport Service	Juice Bar	Indian Restaurant	Snack Place	Food Truck	Fast Food Restaurant	Smoke Shop	Airport Food Court
4	Akurdi	Snack Place	Convenience Store	Indian Restaurant	Fast Food Restaurant	Restaurant	Dumpling Restaurant	Food	Flea Market

# K-mean Cluster Pune

Using K-mean to clustering data area with most common venue

Postcode	District	Area	Taluka	PostalStatus	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	
0	411001	Pune	C D A O	Pune City	Sub Office	18.515806	73.927164	1	Coffee Shop	Indian Restaurant	BBQ Joint	Women's Store	Electronics Store	Food & Drink Shop	Food	Flea Market
1	411001	Pune	Dr B A Chowk	Pune City	Sub Office	18.523306	73.869719	1	Seafood Restaurant	Comfort Food Restaurant	Bed & Breakfast	Sandwich Place	Bakery	Snack Place	Falafel Restaurant	Food
2	411001	Pune	Ghorpuri Bazar	Pune City	Sub Office	18.525956	73.895290	2	Park	Women's Store	Concert Hall	Cupcake Shop	Department Store	Dessert Shop	Diner	Donut Shop
3	411001	Pune	N W College	Pune City	Sub Office	18.518135	73.855552	2	Indian Restaurant	Historic Site	Miscellaneous Shop	Department Store	Ice Cream Shop	Bakery	Falafel Restaurant	Food Court
4	411001	Pune	Pune Cantt East	Pune City	Sub Office	18.457955	73.791910	2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN



# Result

## Cluster 1 for Mumbai

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

	Area	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
16	Asvini	18.900152	72.815235	0	American Restaurant	Bus Station	Fast Food Restaurant	Yoga Studio	Event Space	Food	Flower Shop	Flea Market
17	Colaba	18.906703	72.814712	0	Convenience Store	Bar	Asian Restaurant	Event Space	Food	Flower Shop	Flea Market	Fish Market
41	Jacob Circle	18.982146	72.830131	0	Women's Store	Indian Restaurant	Restaurant	Jewelry Store	Bakery	Yoga Studio	Event Space	Flower Shop
48	Dadar Colony	19.018913	72.852930	0	Deli / Bodega	Coffee Shop	Garden	Café	Yoga Studio	Event Space	Flower Shop	Flea Market
59	Worli Naka	18.999054	72.816961	0	Ice Cream Shop	Indian Restaurant	Seafood Restaurant	Asian Restaurant	Fast Food Restaurant	Café	Bar	Bakery
68	Sion	18.039021	72.861895	0	Indian Restaurant	Gym / Fitness Center	Vegetarian / Vegan Restaurant	Plaza	Bank	Bakery	Café	Movie Theater
69	Nehru Nagar Mumbai	19.063990	72.882563	0	Convenience Store	Cricket Ground	Playground	Arts & Crafts Store	Diner	Chinese Restaurant	Event Space	Flower Shop
72	Cumballa Hill	20.593684	78.962880	0	Spa	Yoga Studio	Electronics Store	Flower Shop	Flea Market	Fish Market	Field	Fast Food Restaurant
79	Gokhale Road Mumbai	19.016786	72.835146	0	Convenience Store	Electronics Store	Indian Restaurant	Fast Food Restaurant	Coffee Shop	Chinese Restaurant	Seafood Restaurant	Flower Shop
98	Reay Road	18.971022	72.845377	0	Bar	Government Building	Yoga Studio	Electronics Store	Flower Shop	Flea Market	Fish Market	Field
103	B P T Colony	19.012526	72.862625	0	Women's Store	Food Court	Flower Shop	Flea Market	Fish Market	Field	Fast Food Restaurant	Farmers Market
109	Government Colony	19.060574	72.850440	0	Indian Restaurant	Bar	Bistro	Café	Restaurant	Multiplex	Pizza Place	Chinese Restaurant
144	Sahar P T Colony	19.100554	72.867843	0	Indian Restaurant	Hotel	Hotel Bar	Asian Restaurant	Coffee Shop	Bar	Yoga Studio	Event Space

## Cluster 2 for Mumbai

```
#Cluster 2 for Mumbai
bintang_merged.loc[bintang_merged['Cluster Labels'] == 1, bintang_merged.columns[[2] + list(range(5, bintang_merged.shape[1]))]]
```

	Area	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
0	Bazargate	18.936518	72.835823	1	Indian Restaurant	Café	Seafood Restaurant	Chinese Restaurant	Irani Cafe	Hostel	Bakery	Park
1	M P T	18.948288	72.844163	1	Harbor / Marina	Hotel	Convenience Store	Cupcake Shop	Cricket Ground	Food	Flower Shop	Flea Market
2	Stock Exchange	18.930012	72.833639	1	Indian Restaurant	Café	Seafood Restaurant	Bookstore	Dessert Shop	Chinese Restaurant	Coffee Shop	Clothing Store
3	Tajmahal	18.921729	72.833031	1	Café	Bar	Hotel	Coffee Shop	Indian Restaurant	Cocktail Bar	Pub	Mughlai Restaurant
4	Town Hall Mumbai	19.115063	72.850029	1	Indian Restaurant	Hotel	Food Court	Burger Joint	Pizza Place	Platform	Café	Chinese Restaurant
5	Mumbai G P O	18.938956	72.837057	1	Indian Restaurant	Café	Irani Cafe	Hostel	Bar	Multiplex	Coffee Shop	Chinese Restaurant
6	Kalbadavi	18.948738	72.828883	1	Indian Restaurant	Jewelry Store	Movie Theater	Bar	Café	Food	Cheese Shop	Yoga Studio
7	Ramwadi	18.950313	72.827458	1	Indian Restaurant	Jewelry Store	Food	Bus Station	Fast Food Restaurant	Event Space	Flower Shop	Flea Market
8	S C Court	18.931120	72.830150	1	Indian Restaurant	Fast Food Restaurant	Café	Coffee Shop	Bakery	Movie Theater	Clothing Store	Chinese Restaurant
9	Thakurdar	18.949847	72.822745	1	Food Truck	Train Station	Aquarium	Fast Food Restaurant	Restaurant	Gastropub	Juice Bar	Electronics Store
10	Ambewadi Mumbai	18.955627	72.821715	1	Juice Bar	Fast Food Restaurant	Breakfast Spot	Vegetarian / Vegan Restaurant	Gym / Fitness Center	Bus Station	Snack Place	Coffee Shop
12	Chauapati	18.954749	72.813882	1	Indian Restaurant	Fast Food Restaurant	Ice Cream Shop	Juice Bar	Café	Pizza Place	Breakfast Spot	Train Station
13	Girgaon	18.957229	72.819669	1	Juice Bar	Fast Food Restaurant	Electronics Store	Indian Restaurant	Coffee Shop	Breakfast Spot	Snack Place	Café
14	Madhavbaug	19.112474	72.848127	1	Indian Restaurant	Hotel	Pizza Place	Yoga Studio	Duty-free Shop	Flea Market	Fish Market	Field
15	Opera House	18.956988	72.815216	1	Fast Food Restaurant	Indian Restaurant	Ice Cream Shop	Coffee Shop	Juice Bar	Electronics Store	Italian Restaurant	Pizza Place
18	Holiday Camp	19.093205	72.913813	1	Café	Vegetarian / Vegan Restaurant	Indian Restaurant	Diner	Lounge	Yoga Studio	Flea Market	Fish Market
19	V W T C	18.914169	72.818242	1	Indian Restaurant	Italian Restaurant	Thai Restaurant	Bar	Spa	Coffee Shop	Food Truck	Shopping Mall
20	Malabar Hill	18.954274	72.801325	1	Park	Gym	Dessert Shop	Restaurant	Electronics Store	Flea Market	Fish Market	Field
140	Seepz	19.126657	72.876655	1	Restaurant	Indian Restaurant	Lake	Airport Terminal	Vegetarian / Vegan Restaurant	Pizza Place	Bus Station	Café
141	Vidyanagari	19.074447	72.864102	1	Fast Food Restaurant	Ice Cream Shop	Indian Restaurant	Pizza Place	Pet Store	Bus Station	Chinese Restaurant	Snack Place
142	Airport Mumbai	19.089560	72.865614	1	Airport	Duty-free Shop	Event Space	Food	Flower Shop	Flea Market	Fish Market	Field
143	International Airport	19.089560	72.865614	1	Airport	Duty-free Shop	Event Space	Food	Flower Shop	Flea Market	Fish Market	Field
145	Sahargaoon	20.593684	78.962880	1	Spa	Yoga Studio	Electronics Store	Flower Shop	Flea Market	Fish Market	Field	Fast Food Restaurant

124 rows × 12 columns



Cluster 3 for Mumbai

#Cluster 3 for Mumbai  
bintang\_merged.loc[bintang\_merged['Cluster Labels'] == 2, bintang\_merged.columns[[2] + list(range(5, bintang\_merged.shape[1]))]]

	Area	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
11	Chamli Road	18.951754	72.818369	2	Juice Bar	Café	Restaurant	Fast Food Restaurant	Train Station	Indian Restaurant	Italian Restaurant	Breakfast Spot

Cluster 1 for Pune

#Cluster 1 for Pune  
jdt\_merged.loc[jdt\_merged['Cluster Labels'] == 0, jdt\_merged.columns[[2] + list(range(5, jdt\_merged.shape[1]))]]

	Area	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
5	Pune	18.457955	73.791910	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
10	Nana Peth	18.520430	73.856744	0	Historic Site	Department Store	Bakery	Women's Store	Falafel Restaurant	Food & Drink Shop	Food	Flea Market
14	Ammunition Factory Khadi	18.564094	73.860245	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
17	Khadi Bazar	18.560054	73.850069	0	Department Store	Medical Supply Store	Bakery	Women's Store	Falafel Restaurant	Food & Drink Shop	Food	Flea Market
21	Film Institute	18.512991	73.828631	0	Café	Snack Place	Seafood Restaurant	General Entertainment	Department Store	Dessert Shop	Sandwich Place	Burger Joint
25	Venwada	18.552878	73.879637	0	Donut Shop	Women's Store	Hotel Bar	Convenience Store	Cupcake Shop	Department Store	Dessert Shop	Diner
42	Dukirkline	18.567315	73.926558	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
43	Vadgaon Sheri	18.552719	73.920817	0	Rajasthan Restaurant	Soccer Field	Women's Store	Food & Drink Shop	Convenience Store	Cupcake Shop	Department Store	Dessert Shop
44	Dhanori	18.583500	73.886488	0	Bus Station	Women's Store	Concert Hall	Cupcake Shop	Department Store	Dessert Shop	Diner	Donut Shop
53	Mazulkar Colony	18.633792	73.814336	0	Dessert Shop	Pizza Place	Chinese Restaurant	Park	Tea Room	Women's Store	Dumpling Restaurant	Flea Market
54	Nehrurang Pune	18.632689	73.822057	0	Hockey Field	Indian Restaurant	Chinese Restaurant	Restaurant	Women's Store	Dumpling Restaurant	Flea Market	Fast Food Restaurant
56	Chinchwad East	18.629781	73.799709	0	Indian Restaurant	Clothing Store	Shopping Mall	Pizza Place	Jewelry Store	BBQ Joint	Chinese Restaurant	Gym
59	Range Hills	18.520430	73.856744	0	Historic Site	Department Store	Bakery	Women's Store	Falafel Restaurant	Food & Drink Shop	Food	Flea Market
61	Sprf	18.491230	73.909528	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
63	Khadawale R S	18.445183	73.785595	0	Snack Place	Indian Restaurant	Women's Store	Dumpling Restaurant	Food	Flea Market	Fast Food Restaurant	Falafel Restaurant
68	Kunal Icon Pimple Saudagar	18.593397	73.795808	0	Café	Sandwich Place	Vegetarian / Vegan Restaurant	Pizza Place	Ice Cream Shop	Department Store	Fast Food Restaurant	Flea Market
69	Roseland Residency Pimple Saudagar	18.593095	73.794667	0	Ice Cream Shop	Pizza Place	Department Store	Fast Food Restaurant	Sandwich Place	Salon / Barbershop	Café	Vegetarian / Vegan Restaurant
71	Sasane Nagar	18.491955	73.934512	0	Department Store	Pizza Place	Café	Fast Food Restaurant	Breakfast Spot	Women's Store	Food Court	Dessert Shop

Cluster 2 for Pune

#Cluster 2 for Pune  
jdt\_merged.loc[jdt\_merged['Cluster Labels'] == 1, jdt\_merged.columns[[2] + list(range(5, jdt\_merged.shape[1]))]]

	Area	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
0	C D A O	18.515806	73.827164	1	Coffee Shop	Indian Restaurant	BBQ Joint	Women's Store	Electronics Store	Food & Drink Shop	Food	Flea Market
1	Dr B A Chowk	18.525306	73.869719	1	Seafood Restaurant	Comfort Food Restaurant	Bed & Breakfast	Sandwich Place	Bakery	Snack Place	Falafel Restaurant	Food
6	Pune New Bazar	18.531229	73.811128	1	Bus Station	Women's Store	Concert Hall	Cupcake Shop	Department Store	Dessert Shop	Diner	Donut Shop
7	Sachapir Street	18.514758	73.876871	1	Indian Restaurant	Fast Food Restaurant	Breakfast Spot	Chinese Restaurant	Snack Place	Ice Cream Shop	Lounge	Indian Chinese Restaurant
46	Yshrantwadi	18.572605	73.876208	1	Indian Restaurant	Pizza Place	South Indian Restaurant	Gym	Café	Breakfast Spot	Dessert Shop	Diner
49	Shivaji Housing Society	18.531617	73.827439	1	Coffee Shop	Lounge	IT Services	Bistro	Nightclub	Pizza Place	Pool Hall	Chinese Restaurant
62	N D A Khadakawale	18.461098	73.759112	1	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
65	Indrapainagar	18.638387	73.845780	1	Convenience Store	Gym	Indian Restaurant	Pizza Place	Food	Flea Market	Fast Food Restaurant	Falafel Restaurant
74	Rashtra Bhawan Bhaon	18.515344	73.849752	1	Indian Restaurant	Ice Cream Shop	Women's Store	Tea Room	Seafood Restaurant	Road	Restaurant	Bar
78	C M E	18.518430	73.931855	1	Indian Restaurant	Coffee Shop	Chinese Restaurant	Italian Restaurant	Donut Shop	Sandwich Place	Multiplex	Bar
85	Thergaon	18.614113	73.772681	1	Accessories Store	Vegetarian / Vegan Restaurant	Mobile Phone Shop	Bed & Breakfast	North Indian Restaurant	Electronics Store	Food	Flea Market
107	P C N T	18.657132	73.765907	1	Indian Restaurant	Snack Place	Bakery	Asian Restaurant	Diner	Women's Store	Food & Drink Shop	Food
109	N I A	18.520430	73.856744	1	Historic Site	Department Store	Bakery	Women's Store	Falafel Restaurant	Food & Drink Shop	Food	Flea Market
110	Katraj	18.457532	73.867746	1	Indian Restaurant	Snack Place	Women's Store	Dumpling Restaurant	Food	Flea Market	Fast Food Restaurant	Falafel Restaurant
111	Lohagaon	18.591129	73.918820	1	Hotel	Hotel Bar	Convenience Store	Cupcake Shop	Department Store	Dessert Shop	Diner	Donut Shop
116	Anandnagar Pune	18.477797	73.821321	1	Indian Restaurant	Snack Place	Pizza Place	Fast Food Restaurant	Bakery	Electronics Store	Food & Drink Shop	Food

Cluster 3 for Pune

#Cluster 3 for Pune  
jdt\_merged.loc[jdt\_merged['Cluster Labels'] == 2, jdt\_merged.columns[[2] + list(range(5, jdt\_merged.shape[1]))]]

	Area	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
2	Ghorpuri Bazar	18.525956	73.895290	2	Park	Women's Store	Concert Hall	Cupcake Shop	Department Store	Dessert Shop	Diner	Donut Shop
3	N W College	18.518135	73.855552	2	Indian Restaurant	Historic Site	Miscellaneous Shop	Department Store	Ice Cream Shop	Bakery	Falafel Restaurant	Food Court
4	Pune Carrit East	18.457955	73.791910	2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
8	Bajirao Road	18.513342	73.853843	2	Indian Restaurant	Dessert Shop	Vegetarian / Vegan Restaurant	Ice Cream Shop	Jewelry Store	Bakery	Women's Store	Snack Place
9	Kapad Ganj	18.514464	73.860143	2	Flea Market	Indian Restaurant	Arts & Crafts Store	Fruit & Vegetable Store	Neighborhood	Women's Store	Electronics Store	Food
11	Pune City	18.457955	73.791910	2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
12	Rainwar Peth	18.517454	73.861241	2	Brewery	Flea Market	Indian Restaurant	Neighborhood	Restaurant	Juice Bar	Food & Drink Shop	Electronics Store
13	Shukrawar Peth Pune	18.508320	73.856700	2	Vegetarian / Vegan Restaurant	Dessert Shop	Jewelry Store	Mobile Phone Shop	History Museum	Indian Restaurant	Tea Room	Breakfast Spot
15	East Khadi	18.569936	73.850643	2	Department Store	Medical Supply Store	Women's Store	Food Truck	Cupcake Shop	Dessert Shop	Diner	Donut Shop
16	H E Factory	18.494207	73.914074	2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
18	Khadi	18.569936	73.850643	2	Department Store	Medical Supply Store	Women's Store	Food Truck	Cupcake Shop	Dessert Shop	Diner	Donut Shop
19	A R Shala	18.520430	73.856744	2	Historic Site	Department Store	Bakery	Women's Store	Falafel Restaurant	Food & Drink Shop	Food	Flea Market
20	Deccan Gymkhana	18.517357	73.841660	2	Indian Restaurant	Vegetarian / Vegan Restaurant	Fast Food Restaurant	Chinese Restaurant	Café	Hotel	Multiplex	Coffee Shop
22	Congress House Road	18.522767	73.853129	2	Indian Restaurant	Bar	Motorcycle Shop	Historic Site	Fast Food Restaurant	Theater	Tea Room	Multiplex
23	S S C Exam Board	18.520430	73.856744	2	Historic Site	Department Store	Bakery	Women's Store	Falafel Restaurant	Food & Drink Shop	Food	Flea Market
24	Shivajinagar Pune	18.530822	73.847465	2	Indian Restaurant	Hotel	Fast Food Restaurant	Pizza Place	Restaurant	Donut Shop	Flea Market	Falafel Restaurant
26	Venwada T S	18.520430	73.856744	2	Historic Site	Department Store	Bakery	Women's Store	Falafel Restaurant	Food & Drink Shop	Food	Flea Market

## Discussion

Based on cluster for each city above, we believe that classification for each cluster can be done better with calculation of venues categories (most common) in each cities. Referring to each cluster, we can't determine clearly what represent in each cluster by using Foursquare - Most Common Venue data.

However, for the sake of this project we assumed each cluster as follow:

**Cluster 1:** Mumbai: Shopping

**Cluster 2:** Mumbai: Mix

**Cluster 3:** Mumbai: Food

**Cluster 1:** Pune: Tourism

**Cluster 2:** Pune: Residential

**Cluster 3:** Pune: Mix

What is lacking at this point is a systematic, quantitative way to identify and distinguish different district and to describe the correlation most common venues as recorded in Foursquare. The reality is however more complex: similar cities might have or might not have similar common venues. A further step in this classification would be to find a method to extract these common venues and integrate the spatial correlations between different of areas or district.

We believe that the classification we propose is an encouraging step towards a quantitative and systematic comparison of the different cities. Further studies are indeed needed in order to relate the data acquired, then observe it to more meaningful and objective results.

## Conclusion

Using Foursquare API, we can captured data of common places all around the world. Using it, we refer back to our main objectives, which is to determine;

the similarity or dissimilarity of both cities classification of area located inside the city whether it is residential, tourism places, or others In conclusion, both cities Mumbai and Pune are the center of attraction among Indian. However, to declare both cities are similar or dissimilar base on common venues visited is quite difficult. Both cities is similar in some venues also dissimilar in certain venues. And for classification based on common venues, again we must have more systematic or quantitative way to identify and declare this. Comparison can be made, but no such method or quantitative data to determine this. We hope in the future, a method to determine it can be establish and explore for references.