

The Battle of Neighbourhoods

Finding the best area to invest on a luxury hotel in **Gurgaon City, India**

Date: 18-11-2019

Author: Deepak Kumar Ojha

1. Introduction

1.1. Background

This comprehensive project report is for anyone who is seeking data facts about opening a new luxury hotel in the city of Gurgaon, India. It provides a clear recommendation on what would be the best venue to start a new hotel in a highly visited city with already many good hotels.

Gurgaon, officially named Gurugram, is a city located in the northern Indian state of Haryana. It is situated near the Delhi-Haryana border, about 30 kilometres (19 mi) southwest of the national capital New Delhi and 268 km (167 mi) south of Chandigarh, the state capital. It is one of the major satellite cities of Delhi and is part of the National Capital Region of India. As of 2011, Gurgaon had a population of 876,900. Gurgaon has become a leading financial and industrial hub with the third-highest per capita income in India. The city's economic growth story started when the leading Indian automobile manufacturer Maruti Suzuki India Limited established a manufacturing plant in Gurgaon in the 1970s. Today, Gurgaon has local offices for more than 250 Fortune 500 companies. Gurgaon is categorised as very high on the Human Development Index, with an HDI of 0.889 (2017), which is also the highest in India.

1.2. Business Problem

This report focusses on the issue of deciding best neighbourhood to open a new hotel in a city like Gurgaon. Let's imagine that my client company is willing to open a new luxury hotel, and first and foremost important decision will be the location for the hotel.

Below are some questions which will help solving the business problem.

1. What are the highest rated hotel locations in Gurgaon city?
2. Which place has a greater number of restaurants in proximity?
3. Which neighbourhood is the best location for opening a luxury hotel?

2. Data Section

For this capstone project, we will be using 2 data sources for our analysis.

Data Source 1: Zomato Dataset (from Kaggle)

Description: This dataset contains Locality data of Gurgaon, India along with their latitude and longitude

Data Source 2: Foursquare API

Description: Using this API, we will fetch the neighbourhoods of each Locality of Gurgaon, India

3. Methodology Section

The Methodology section comprises of below stages:

1. Collect Data and load into pandas dataframe
2. Data Cleaning & Pre-processing
3. Exploratory Data Analysis
4. Use Foursquare API to get neighbourhood venues
5. Visualize the locality using folium library
6. Modelling using k-means clustering based on common neighbourhoods
7. Derive conclusions from the model by analysing each cluster

3.1. Data Collection & Loading

Zomato locality data is downloaded from Kaggle website which contains locality name, restaurant names, ratings, cuisines, latitude, longitude etc.

Data is then filtered to take only for Gurgaon City, India.

```
df['City'].value_counts().to_frame().head()
```

	City
New Delhi	5473
Gurgaon	1118
Noida	1080
Faridabad	251
Ghaziabad	25

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longitude	Latitude	Cuisines	...	Currency	Has Table booking	Has Online delivery
0	18396451	K Lab	1	Gurgaon	Shop GF-18, ILD Trade Centre, Sector 47, Near ...	ILD Trade Centre Mall, Sohna Road	ILD Trade Centre Mall, Sohna Road, Gurgaon	77.039310	28.424831	Cafe, Beverages	...	Indian Rupees(Rs.)	No	No
1	18237941	Pind Balluchi	1	Gurgaon	112/112-A, 1st Floor, ILD Trade Centre, Near S...	ILD Trade Centre Mall, Sohna Road	ILD Trade Centre Mall, Sohna Road, Gurgaon	77.039220	28.424913	North Indian, Mughlai	...	Indian Rupees(Rs.)	Yes	Yes

3.2. Data Cleaning & Pre-Processing

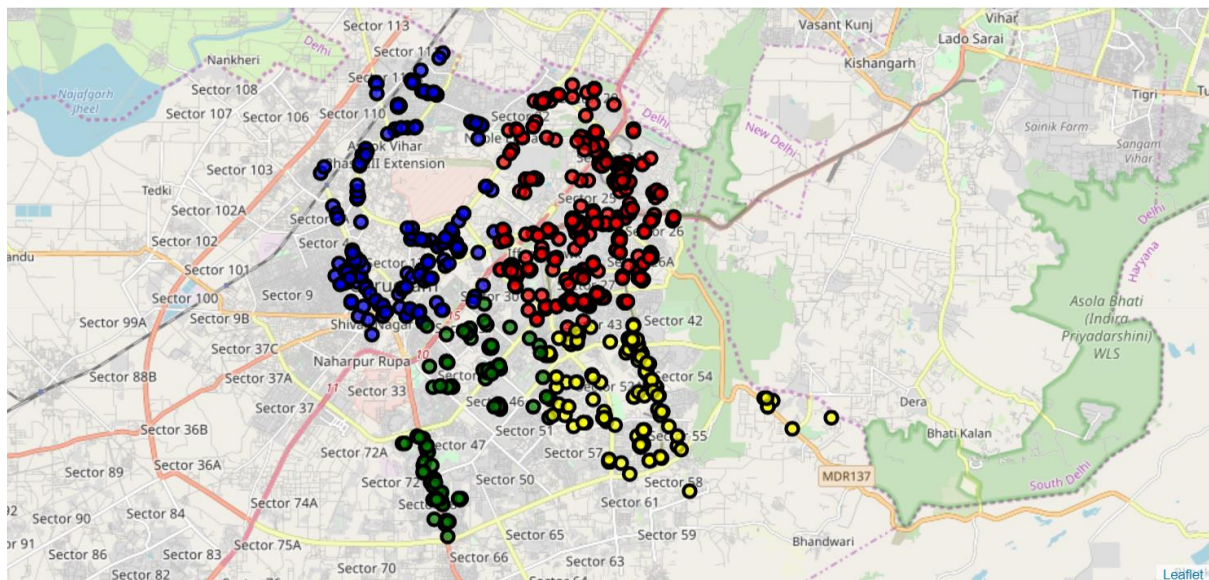
Meaningful column names are included in the dataset. Only relevant columns are kept, rest are discarded. Final data frame is as below;

```
dfClean.head()
```

	Restaurant Name	Locality	Longitude	Latitude	Cuisines	Aggregate rating	Rating text	Votes
0	K Lab	ILD Trade Centre Mall, Sohna Road	77.039310	28.424831	Cafe, Beverages	3.4	Average	16
1	Pind Balluchi	ILD Trade Centre Mall, Sohna Road	77.039220	28.424913	North Indian, Mughlai	2.7	Average	80
2	Punjab Grill	Ambience Mall, Gurgaon	77.097118	28.503077	North Indian, Mughlai	4.3	Very Good	1887
3	Zambar	Ambience Mall, Gurgaon	77.097185	28.502613	South Indian, Seafood, Kerala	4.0	Very Good	802
4	Cakes & More	Ansai Plaza Mall, Palam Vihar	77.042009	28.511416	Bakery	3.0	Average	4

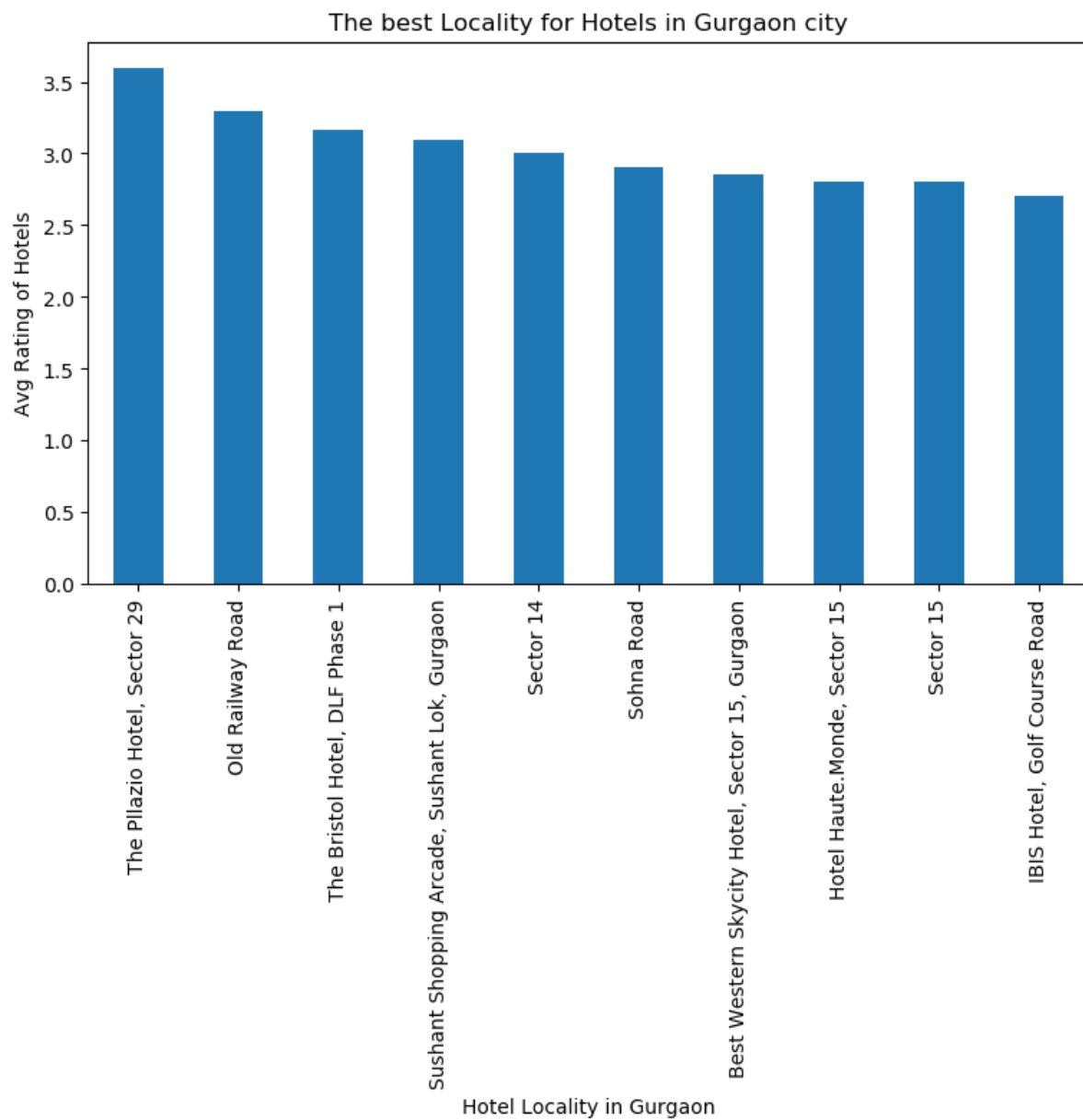
3.3. Data Visualization using Folium library

The locations are plotted into the map using python folium library to visualize the spread.



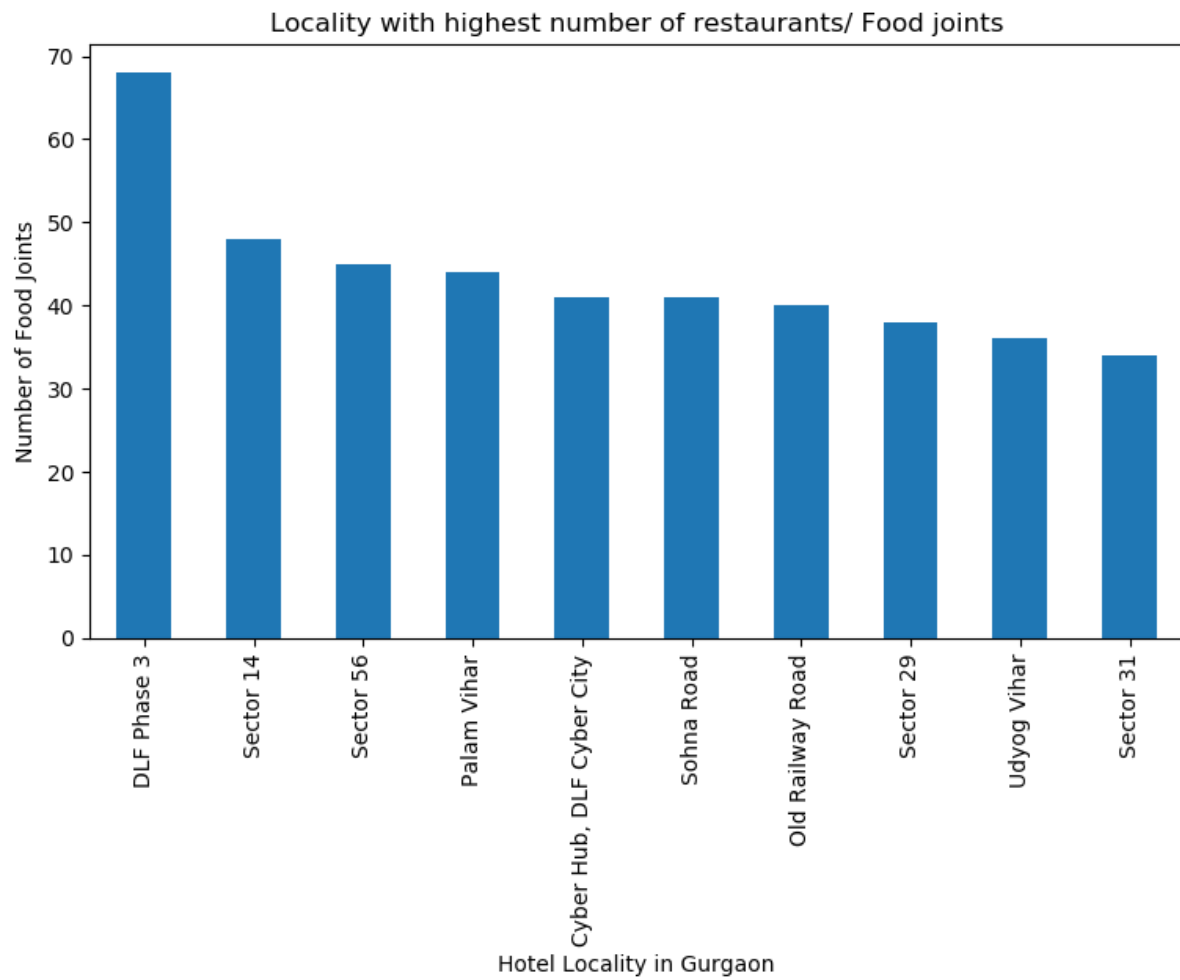
3.4. Data Analysis

The data is analysed to find out the best location for highest rated hotels already situated in Gurgaon.



As per the above graphical representation, it is clear that **Sector-29, Gurgaon** is the locality for hotels having highest ratings by customers.

Then, the data is analysed to find out the location where highest number of restaurants or food joints are situated in Gurgaon.



DLF Phase-3 has highest number of restaurants/food joints.

3.5. Foursquare API (To get venues in each neighbourhood)

Foursquare API is used to find all venues in locality of Gurgaon.

```
ggn_venues.head()
```

	Locality	Locality Latitude	Locality Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	ILD Trade Centre Mall, Sohna Road	28.424872	77.039265	Brewer Street	28.424499	77.039134	Brewery
1	ILD Trade Centre Mall, Sohna Road	28.424872	77.039265	Decathlon Gurgaon	28.427433	77.036258	Sporting Goods Shop
2	ILD Trade Centre Mall, Sohna Road	28.424872	77.039265	KFC	28.424789	77.038722	Fast Food Restaurant
3	ILD Trade Centre Mall, Sohna Road	28.424872	77.039265	Sagar Ratna	28.421042	77.038509	Indian Restaurant
4	ILD Trade Centre Mall, Sohna Road	28.424872	77.039265	Grand cinemas	28.423626	77.039422	Movie Theater

```
ggn_venues.groupby('Locality').count()
```

	Locality Latitude	Locality Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Locality						
ILD Trade Centre Mall, Sohna Road	11	11	11	11	11	11
Ambience Mall, Gurgaon	38	38	38	38	38	38
Ansai Plaza Mall, Palam Vihar	4	4	4	4	4	4
Best Western Skycity Hotel, Sector 15, Gurgaon	5	5	5	5	5	5
Central Arcade, DLF Phase 2, Gurgaon	21	21	21	21	21	21

The Habitare Hotel, Sector 14	5	5	5	5	5	5
The Oberoi, Udyog Vihar	13	13	13	13	13	13
The Pilazio Hotel, Sector 29	30	30	30	30	30	30
The Sapphire Mall, Sohna Road	4	4	4	4	4	4
The Westin Gurgaon, Sector 29	7	7	7	7	7	7
The Westin Sohna Resort & Spa, Sohna Road	1	1	1	1	1	1
Trident, Udyog Vihar	13	13	13	13	13	13
Two Horizon Center, Golf Course Road	26	26	26	26	26	26
Unitech Cyber Park, Sector 39, Gurgaon	12	12	12	12	12	12
Unitech Infospace, Sector 21, Gurgaon	10	10	10	10	10	10
Vivanta by Taj, Sector 44, Gurgaon	13	13	13	13	13	13
Vyapar Kendra, Palam Vihar	4	4	4	4	4	4
Vyapar Kendra, Sushant Lok	18	18	18	18	18	18

```
print('There are {} uniques categories.'.format(len(ggn_venues['Venue Category'].unique())))
```

There are 137 uniques categories.

3.6. One-Hot Encoding

In first step of cluster analysis, one hot encoding was conducted to give binary values to each venue category

	Locality	American Restaurant	Arcade	Art Gallery	Asian Restaurant	Athletics & Sports	BBQ Joint	Badminton Court	Bagel Shop	Bakery	...	Thai Restaurant	Theater	Tibetan Restaurant	Trail	Train Station	Vegetarian / Vegan Restaurant	Whisky Bar
0	ILD Trade Centre Mall, Sohna Road	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
1	ILD Trade Centre Mall, Sohna Road	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
2	ILD Trade Centre Mall, Sohna Road	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
3	ILD Trade Centre Mall.	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0


```
ggn_grouped = ggn_onehot.groupby('Locality').mean().reset_index()
ggn_grouped
```



```
]:
```

	Locality	American Restaurant	Arcade	Art Gallery	Asian Restaurant	Athletics & Sports	BBQ Joint	Badminton Court	Bagel Shop	Bakery	...	Thai Restaurant	Theater	Tibetan Restaurant	Trail	Train Station
0	ILD Trade Centre Mall, Sohna Road	0.000000	0.000000	0.000000	0.000000	0.000000	0.090909	0.000000	0.000000	0.000000	...	0.000000	0.0	0.0	0.0	0.000000
1	Ambience Mall, Gurgaon	0.052632	0.026316	0.000000	0.026316	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.0	0.0	0.0	0.000000
2	Ansal Plaza Mall, Palam Vihar	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.0	0.0	0.0	0.000000
3	Best Western Skycity Hotel, Sector 15, Gurgaon	0.000000	0.000000	0.000000	0.200000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.0	0.0	0.0	0.000000
13	Cyber Greens, DLF Cyber City	0.016667	0.000000	0.000000	0.050000	0.016667	0.016667	0.000000	0.000000	0.000000	...	0.000000	0.0	0.0	0.0	0.000000
14	Cyber Hub, DLF Cyber City	0.016667	0.000000	0.000000	0.050000	0.016667	0.016667	0.000000	0.000000	0.000000	...	0.000000	0.0	0.0	0.0	0.000000
15	DLF Cyber City	0.016129	0.000000	0.000000	0.032258	0.016129	0.016129	0.000000	0.000000	0.000000	...	0.000000	0.0	0.0	0.0	0.000000
16	DLF Galleria, DLF Phase 4	0.031250	0.000000	0.000000	0.031250	0.000000	0.000000	0.000000	0.031250	0.093750	...	0.000000	0.0	0.0	0.0	0.000000
17	DLF Mega Mall, DLF Phase 1	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.0	0.0	0.0	0.000000

Then, the data was grouped by locality names to find out how many venues of each category exist in the locations. We attained a list of 10 most common venue categories in each locality as follows.

---- ILD Trade Centre Mall, Sohna Road----

	venue	freq
0	Movie Theater	0.18
1	Indian Restaurant	0.18
2	Brewery	0.18
3	Sporting Goods Shop	0.09
4	Fast Food Restaurant	0.09
5	Shopping Mall	0.09
6	Department Store	0.09
7	BBQ Joint	0.09
8	Korean Restaurant	0.00
9	New American Restaurant	0.00

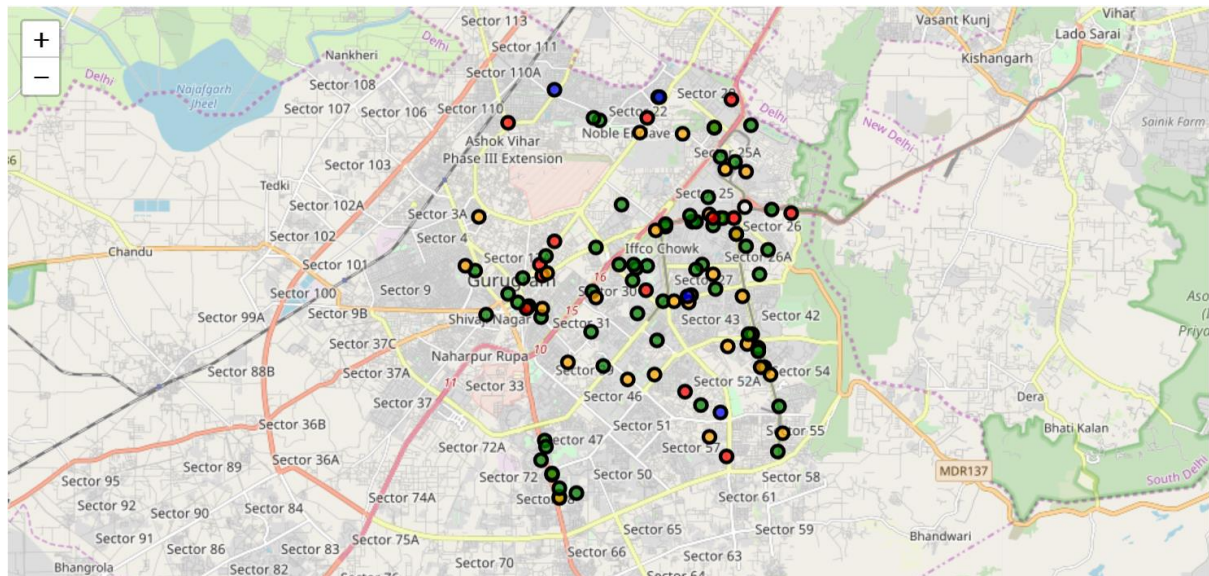
----Ambience Mall, Gurgaon----

	venue	freq
0	Indian Restaurant	0.11
1	Italian Restaurant	0.08
2	Fast Food Restaurant	0.08
3	American Restaurant	0.05
4	Bar	0.05
5	Shopping Mall	0.05
6	Clothing Store	0.05
7	Café	0.05
8	Multiplex	0.05
9	Gym	0.03

	Locality	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	10th Most Common Venue
0	ILD Trade Centre Mall, Sohna Road	Indian Restaurant	Movie Theater	Brewery	Department Store	Shopping Mall	BBQ Joint	Fast Food Restaurant	Sporting Goods Shop	Yoga Studio	Falafel Restaurant
1	Ambience Mall, Gurgaon	Indian Restaurant	Fast Food Restaurant	Italian Restaurant	American Restaurant	Café	Clothing Store	Multiplex	Bar	Shopping Mall	Movie Theater
2	Ansal Plaza Mall, Palam Vihar	Pizza Place	Food	Café	Flea Market	Fast Food Restaurant	Farmers Market	Falafel Restaurant	Electronics Store	Dumpling Restaurant	Yoga Studio
3	Best Western Skycity Hotel, Sector 15, Gurgaon	Hotel	Asian Restaurant	Farmers Market	Dumpling Restaurant	Food	Flea Market	Fast Food Restaurant	Falafel Restaurant	Electronics Store	Donut Shop

3.7. Modelling – k-Means Clustering Algorithm

Based on the venue categories, K-means clustering was conducted to group the neighbourhoods into 5 different clusters based on their similarity. The coloured dots below represent different clusters.



3.8. Analyse each of the clusters

After observing each clusters and the characteristics they possess, we have given names for each clusters that best depicts their characteristics.

Cluster – 0: Best location for Departmental Stores

7	Suites By Carlson, Udyog Vihar	28.509225	77.091835	3.500000	Good	32	0	Clothing Store	Hotel	Fast Food Restaurant	Campground	Cosmetics Shop	Dessert Shop
8	Country Inn & Suites by Carlson, Gurgaon	28.461988	77.067644	3.800000	Good	91	0	Hotel	Japanese Restaurant	Go Kart Track	Indian Restaurant	Resort	Electronics Store
9	Country Inn & Suites by Carlson, Sector 12	28.468385	77.037736	3.250000	Average, Good	44	0	Hotel	Indian Restaurant	Dumpling Restaurant	Food	Flea Market	Fast Food Restaurant
33	Hong Kong Bazaar Mall, Sector 57, Gurgaon	28.421012	77.090534	3.000000	Average	333	0	Convenience Store	Dessert Shop	Shopping Mall	Café	Sandwich Place	Bakery

Cluster – 1: Best location for Luxury Hotel

62	Ramada Gurgaon Central, Sector 44	28.449740	77.070652	3.500000	Good	194	1	Hotel	Gym	Snack Place	Café	Dessert Shop	Dej Sto
63	Sadar Bazar	28.461005	77.029078	3.150000	Average	31	1	Hotel	Market	Fast Food Restaurant	Donut Shop	Electronics Store	Foc
65	Sector 12	28.464958	77.032920	2.000000	Average, Not rated	13	1	Hotel	Dessert Shop	Bus Station	Electronics Store	Food	Fle. Ma
68	Sector 17	28.483299	77.060899	0.200000	Average, Not rated	265	1	Wine Shop	Lounge	Electronics Store	Food	Flea Market	Fas Res
69	Sector 21	28.509766	77.071145	2.116667	Average, Good, Not rated	98	1	Pizza Place	Diner	Fast Food Restaurant	Donut Shop	Sandwich Place	Caf
71	Sector 23	28.503855	77.054534	2.175000	Average, Good, Not rated	332	1	Hotel	Sports Club	Café	Electronics Store	Food	Fle. Ma
72	Sector 29	28.468295	77.064269	3.892105	Average, Excellent, Good, Very	20881	1	Brewery	Indian Restaurant	Café	Hotel	Chinese Restaurant	Fas Res

Cluster – 2: Best location for Café or Bakery

94	Sushant Shopping Arcade, Sushant Lok, Gurgaon	28.460448	77.079375	3.194444	Average, Good, Poor, Very Good	598	2	Hotel	Café	Restaurant	Indian Restaurant	Chinese Restaurant	Fast Food Restaurant
103	The Westin Sohna Resort & Spa, Sohna Road	28.239367	77.142751	3.550000	Average, Good	67	2	Indian Restaurant	Yoga Studio	Dumpling Restaurant	Food	Flea Market	Fast Food Restaurant
108	Unitech Infospace, Sector 21, Gurgaon	28.509796	77.071568	2.950000	Average, Good, Not rated, Poor	973	2	Pizza Place	Diner	Fast Food Restaurant	Donut Shop	Sandwich Place	Café

Cluster – 3: Best location for Hotels

84	Sector 7	28.468027	77.016827	3.240000	Average, Good	431	3	Bakery	Chinese Restaurant	Multiplex	Snack Place	Dessert Shop	Departm Store
93	Sushant Lok	28.459051	77.079674	2.878571	Average, Good, Not rated, Poor, Very Good	1055	3	Hotel	Café	Restaurant	Indian Restaurant	Chinese Restaurant	Fast Foo Restaura
98	The Habitare Hotel, Sector 14	28.466273	77.039830	3.000000	Average	16	3	Hotel	Indian Restaurant	Dumpling Restaurant	Food	Flea Market	Fast Foo Restaura
99	The Oberoi, Udyog Vihar	28.502287	77.086799	3.200000	Average	17	3	Hotel	Restaurant	Seafood Restaurant	Pool	Clothing Store	Indian Restaura
102	The Westin Gurgaon, Sector 29	28.476763	77.070417	3.866667	Good	1056	3	Hotel	Italian Restaurant	Hotel Bar	Fast Food Restaurant	Brewery	Breakfas Spot

Cluster – 4: Best location for Coworking/Small offices/Furniture Shops

	Locality	Latitude	Longitude	Aggregate rating	Rating text	Votes	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
87	Sikandarpur	28.482529	77.095601	0.0	Not rated	0	4	Coworking Space	Furniture / Home Store	Speakeasy	Café	Coffee Shop	Department Store

4. Results & Discussion

We analysed the locality data and neighbourhood data using Foursquare APIs. Using One-hot encoding to get most common venues around localities. Applied k-Means clustering algorithm with 5 clusters and we draw below results;

1. **Sector-29** is the best locality for highest rated hotels
2. **DLF Phase-3** has the highest number of restaurants
3. Hotels are very popular in **Cluster 1 & Cluster 3** as it is within close proximity to restaurants (Italian, Chinese, Continental), Brewery, Coffee shops, Wine Shop, Diner, Sports Club and Shopping malls
4. **Cluster 0** is popular for Departmental Stores
5. **Cluster 2** is popular for Cafe or Bakery
6. **Cluster 4** is popular for Co-working space, small offices or Furniture Stores

The results are expected to be more refined and accurate once we get more datasets.

5. Conclusion

It is evident that from our analysis that **Cluster 1 and Cluster 3** localities are suitable for opening a new luxury hotel. Within close proximity to number of restaurants, coffee shops, shopping malls and breweries, it helps the decision stronger

Sector-29, Sector-14, Sector-15, Udyog Vihar are some of the best locations for opening a Luxury Hotel in the city of Gurgaon, India.