Capstone Project

Battle of Neighborhoods-Paris

To set up new Indian restaurant in Paris

1. Introduction

1.1 Background

Paris is the capital and the most populous city of France and has a population of more than 2.1 million people. Since the 17th century, Paris has been one of Europe's major centres of finance, diplomacy, commerce, fashion, science and arts. It is the second most expensive city in the world so the residents living closer to the city center are likely to have higher

income. And Paris is the most visited city in the world due to its famous heritage sites and museums. With its diverse culture, comes diverse food items. There are many restaurants in Paris, each belonging to different categories like Chinese, Indian , French etc.

1.2 Business Problem

The main objective is to set up a new Indian restaurant in the best location in Paris.If any entrepreneur wishes to set up a new Indian restaurant,where should he/she would like to open in a best suitable place in the neighborhood and have the best chance of being successful.In order to achieve high success rate we must consider following factors:

- Market Places
- Competition in particular location
- Aiding places that make people come to restaurants like Gym, Entertaining Public places
- Population
- Menu from competitors.

Considering these factors, we can find the solution by using data sources provided by Wikipedia and Foursquare API and some Clustering methods.

1.3 Target Audience

- Business personnel who want to invest or open a restaurant.
- Freelancer who loves to have their own restaurant as a side business.
- Finding the best location for opening a restaurant.
- Budding Data Scientists, who want to implement some of the most used Exploratory Data Analysis techniques to obtain necessary data, analyze it and, finally, be able to tell a story out of it.

2.Data Acquisition

The data is acquired from the following sources:

- https://opendata.paris.fr/explore/dataset/arrondissements/table/?dataChart

- ➢ https://data.opendatasoft.com/explore/dataset/arrondissements%40parisdata/export/
- ➢ https://en.wikipedia.org/wiki/Arrondissements_of_Paris

These are the links to Paris data which consists of neighborhood names ,Arrondissement numbers,perimeter,longitude,latitude,and so on. We filter this data according to our requirements.We can also consider the population present in each neighborhood for our data.

We obtain geographical coordinates of Paris city and also the venues in each neighborhood using Foursquare API.We can also filter these venues to get only Indian Restaurants.This can help us in finding the best suitable place for setting up the restaurant.

3.Data Cleaning

We need to outline the initial data that is required:

- District data for Paris including names, location data if available, and any other details required.

Data Analysis and Location Data:

- Foursquare location data will be leveraged to explore or compare districts around Paris.
- Data manipulation and analysis to derive subsets of the initial data.
- Identifying the high traffic areas using data visualisation and statistical analysis.

Visualization:

- Analysis and plotting visualizations.
- Data visualization using various mapping libraries.

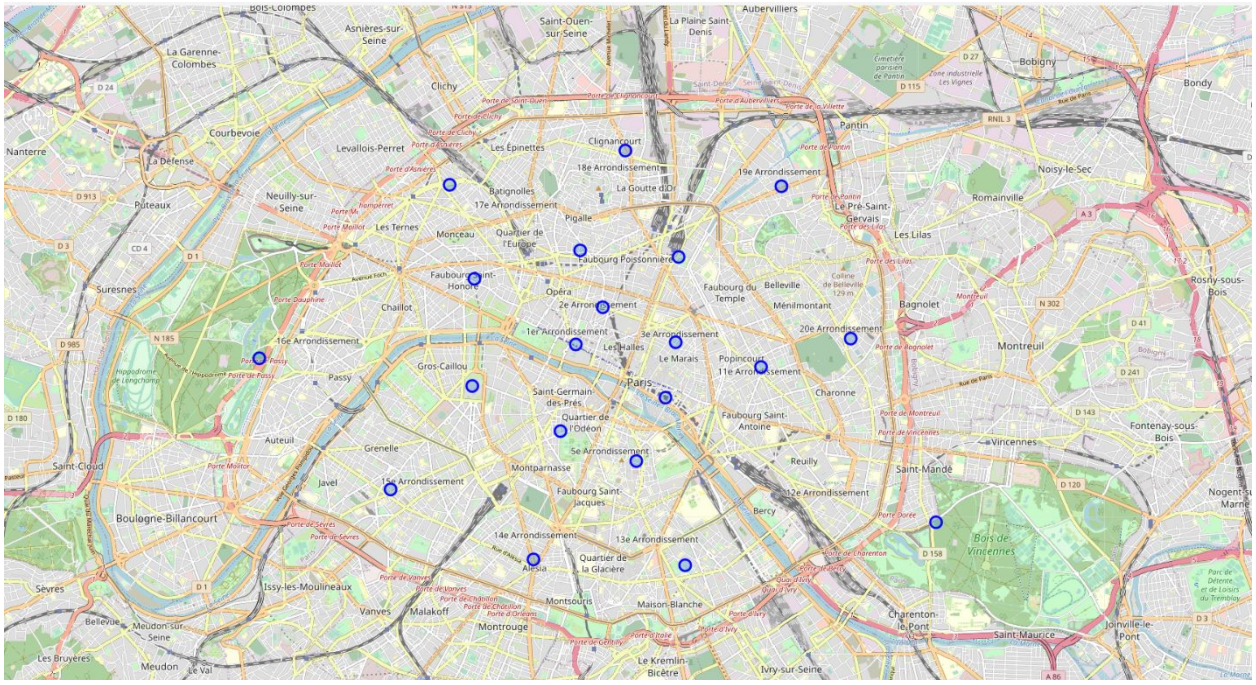
Discussion and Conclusions:

- Recommendations and results based on the data analysis.
- Discussion of any limitations and how the results can be used, and any conclusions that can be drawn.

4.Methodology

4.1 Exploratory Data Analysis

Using the Geopy library and Foursquare API, we can obtain the coordinates of Paris and explore nearby venues of each arrondissement of Paris.We can explore the Arrondissements of Paris using Folium maps.



In this map,we can see the arrondissements are represented in blue circles.There are 20 arrondissements in Paris.

Using the Foursquare API, we can extract nearly 100 venues of each arrondissement in Paris.Using this information,we can create a dataframe that consists of arrondissement, venue names,geolocation data and venue category.Here we can explore the data and perform analysis on it.We can find how many venues are there for each arrondissement and how many unique venue categories are present.

	Neighborhood	Latitude	Longitude	Venue	Venue_Latitude	Venue_Longitude	Venue_Category
0	Temple	48.862872	2.360001	Mmmozza	48.863910	2.360591	Sandwich Place
1	Temple	48.862872	2.360001	Chez Alain Miam Miam	48.862369	2.361950	Sandwich Place
2	Temple	48.862872	2.360001	Square du Temple	48.864475	2.360816	Park
3	Temple	48.862872	2.360001	Marché des Enfants Rouges	48.862806	2.361996	Farmers Market
4	Temple	48.862872	2.360001	Chez Alain Miam Miam	48.862781	2.362064	Sandwich Place
5	Temple	48.862872	2.360001	Okomusu	48.861453	2.360879	Okonomiyaki Restaurant
6	Temple	48.862872	2.360001	Le Burger Fermier des Enfants Rouges	48.862831	2.362073	Burger Joint
7	Temple	48.862872	2.360001	Hôtel Jules & Jim	48.863496	2.357395	Hotel
8	Temple	48.862872	2.360001	SoMa	48.861511	2.362146	Japanese Restaurant
9	Temple	48.862872	2.360001	Les Enfants Rouges	48.863013	2.361260	Wine Bar
10	Temple	48.862872	2.360001	Bontemps	48.863956	2.360725	Dessert Shop
11	Temple	48.862872	2.360001	Biglove Caffè	48.862063	2.363557	Italian Restaurant
12	Temple	48.862872	2.360001	Little Red Door	48.863703	2.363514	Speakeasy
13	Temple	48.862872	2.360001	Hank Burger	48.861340	2.358304	Burger Joint
14	Temple	48.862872	2.360001	Lily of the Valley	48.865221	2.361990	Tea Room
15	Temple	48.862872	2.360001	L'Îlot	48.864666	2.363317	Seafood Restaurant
16	Temple	48.862872	2.360001	Breizh Café	48.860613	2.361804	Creperie
17	Temple	48.862872	2.360001	Candelaria	48.863032	2.364059	Cocktail Bar
18	Temple	48.862872	2.360001	Taing Song-Heng	48.864701	2.356888	Vietnamese Restaurant
19	Temple	48.862872	2.360001	0fr.	48.865746	2.361236	Bookstore
20	Temple	48.862872	2.360001	Le Roi de Pique	48.863375	2.358299	Café
21	Temple	48.862872	2.360001	Hank Pizza	48.863737	2.357199	Pizza Place
22	Temple	48.862872	2.360001	Le Barav	48.865166	2.363155	Wine Bar
23	Temple	48.862872	2.360001	Gigi	48.864753	2.363088	Creperie
24	Temple	48.862872	2.360001	Laïzé 來座 (Laïzé)	48.863077	2.355073	Bubble Tea Shop
25	Temple	48.862872	2.360001	Fringe	48.862612	2.364752	Coffee Shop

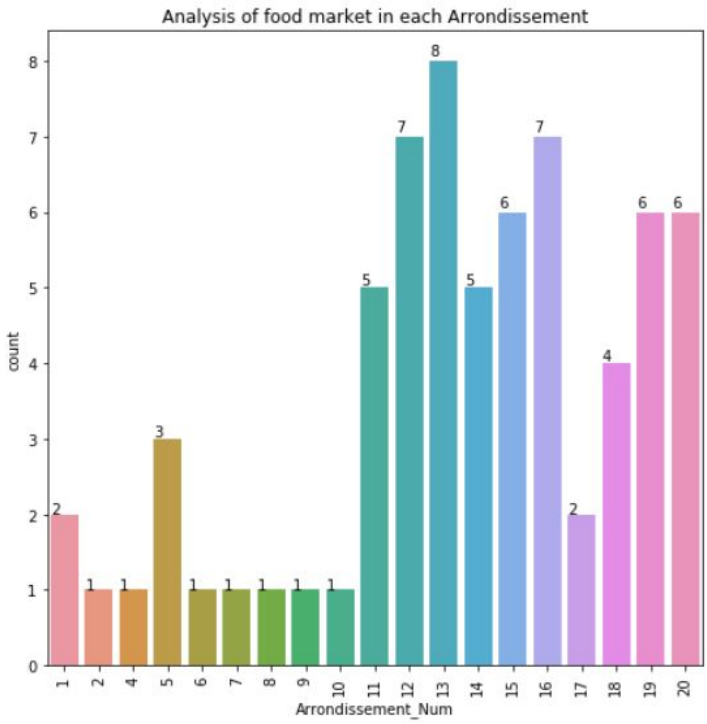
Using onehot encoding we group the arrondissements by mean of the frequency of occurrence of each category.After analysing each arrondissement in Paris,we can create a dataframe in which top 10 common venue categories are sorted for each arrondissement.

Now we can retrieve the farmer's market data from data sources and visualize the obtained data using a seaborn library.The dataset below gives information about existing markets in Paris for buying ingredients. Columns consists on:

- ardt: Arrondissement
- dimanche: if market opens Sunday (1) if not (0)
- geo_point_2d: Localisation of market
- gestionnaire: name of company that manage the market
- Information about opening time and closing time
- id_marche: Identification of market
- localisation: Address of market
- jours_tenue: Days of opening of markets

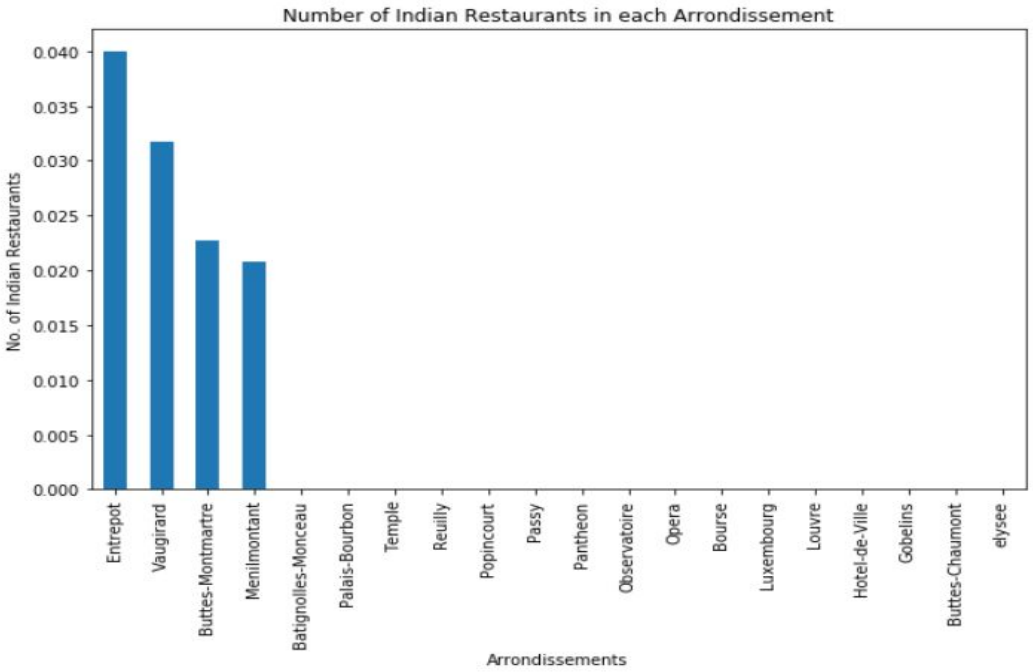
- produit: products that are bought in the market

Here in the plot, we can see that arrondissements 12, 13 and 16 i.e Reuilly,Gobelins,and Passy have the most number of food markets then compared to other arrondissements. So for freshness and lower cost of ingredients, we can consider these arrondissements for setting up the restaurant.

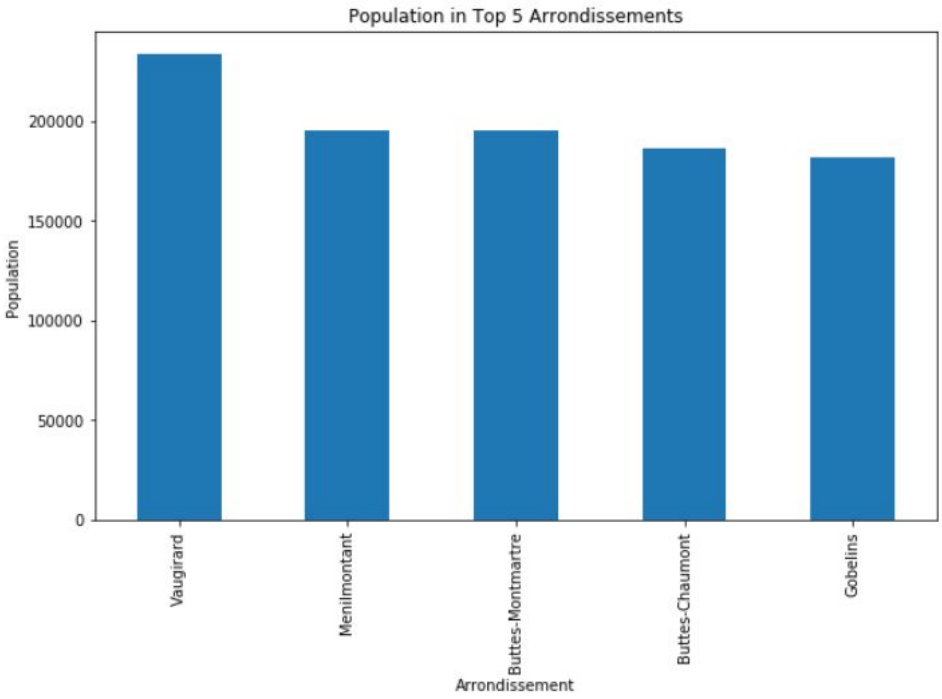


Now using the venue data, we can find the number of Indian Restaurants present in each arrondissement.By visualizing, we can see that only four arrondissements are having Indian restaurants i.e Entrepot,Buttes-Montmartre,Ménilmontant,and Vaugirard.Here the plot shows that Entrepot has most number of Indian restaurants and followed by Vaugirard,Buttes-Montmartre and Ménilmontant.

Now we group the arrondissements by calculating the mean of frequency of occurrence of Indian restaurants.



Now we plot a bar chart which displays the top five highly populated arrondissements using matplotlib library.The plot shows that Vaugirard, Ménilmontant, Buttes-Montmartre, Buttes-Chaumont and Gobelins.



The next part is to group the neighborhoods into clusters using the K-means clustering algorithm.We group the neighborhoods into three clusters which consists of 20 arrondissements of Paris.

Cluster 0 consists of 16 Arrondissements along with the top 10 most common venue categories.Cluster 1 and Cluster 2 consist of two arrondissements each along with the top 10 most common venue categories.

	Arrondissement_Num	Neighborhood	French_Name	Latitude	Longitude	Population_in_2019	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	Cluster Labels
0	3	Temple	3eme Ardt	48.862872	2.360001	34788	French Restaurant	Italian Restaurant	Japanese Restaurant	Art Gallery	Gourmet Shop	Wine Bar	Bakery	Coffee Shop	Burger Joint	Cocktail Bar	0
2	14	Observatoire	14eme Ardt	48.829245	2.326542	137105	French Restaurant	Hotel	Bistro	Bakery	Plaza	Tea Room	Italian Restaurant	Pizza Place	Brasserie	Food & Drink Shop	0
3	10	Entrepot	10eme Ardt	48.876130	2.360728	91932	French Restaurant	Hotel	Coffee Shop	Bistro	Indian Restaurant	Café	Bar	Pizza Place	Japanese Restaurant	Breakfast Spot	0
4	12	Reuilly	12eme Ardt	48.834974	2.421325	141494	Zoo Exhibit	Bike Rental / Bike Share	Supermarket	Monument / Landmark	Zoo	Coworking Space	Donut Shop	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	0
5	16	Passy	16eme Ardt	48.860392	2.261971	165446	Park	Plaza	Lake	Boat or Ferry	Trail	Bus Station	French Restaurant	Pool	Art Museum	Recording Studio	0
6	11	Popincourt	11eme Ardt	48.859059	2.380058	147017	French Restaurant	Café	Pastry Shop	Wine Bar	Restaurant	Supermarket	Bakery	Asian Restaurant	Italian Restaurant	Japanese Restaurant	0
8	4	Hotel-de-Ville	4eme Ardt	48.854341	2.357630	27487	French Restaurant	Ice Cream Shop	Clothing Store	Hotel	Pedestrian Plaza	Italian Restaurant	Plaza	Wine Bar	Gay Bar	Garden	0
11	1	Louvre	1er Ardt	48.862563	2.336443	16252	French Restaurant	Japanese Restaurant	Plaza	Hotel	Café	Art Museum	Italian Restaurant	Garden	Brasserie	Theater	0
12	5	Pantheon	5eme Ardt	48.844443	2.350715	59108	French Restaurant	Hotel	Bakery	Italian Restaurant	Science Museum	Coffee Shop	Bar	Plaza	Pub	Historic Site	0
13	7	Palais-Bourbon	7eme Ardt	48.856174	2.312188	52512	Hotel	French Restaurant	Café	Italian Restaurant	Plaza	Art Museum	History Museum	Coffee Shop	Cocktail Bar	Ice Cream Shop	0
14	20	Menilmontant	20eme Ardt	48.863461	2.401188	195604	Bakery	Japanese Restaurant	Pizza Place	French Restaurant	Plaza	Bar	Café	Supermarket	Italian Restaurant	Bistro	0
15	8	elysee	8eme Ardt	48.872721	2.312554	36453	French Restaurant	Hotel	Corsican Restaurant	Spa	Art Gallery	Theater	Sporting Goods Shop	Furniture / Home Store	Park	Resort	0
16	9	Opera	9eme Ardt	48.877164	2.337458	59629	French Restaurant	Hotel	Cocktail Bar	Bakery	Bistro	Wine Bar	Japanese Restaurant	Lounge	Plaza	Café	0
17	13	Gobelins	13eme Ardt	48.828388	2.362272	181552	Vietnamese Restaurant	Asian Restaurant	Thai Restaurant	Chinese Restaurant	French Restaurant	Juice Bar	Plaza	Cambodian Restaurant	Furniture / Home Store	Cosmetics Shop	0
18	15	Vaugirard	15eme Ardt	48.840085	2.292826	233484	Hotel	French Restaurant	Italian Restaurant	Coffee Shop	Supermarket	Indian Restaurant	Bistro	Japanese Restaurant	Lebanese Restaurant	Brasserie	0
19	6	Luxembourg	6eme Ardt	48.849130	2.332898	40916	French Restaurant	Cocktail Bar	Bakery	Bistro	Fountain	Pastry Shop	Hotel	Plaza	Miscellaneous Shop	Seafood Restaurant	0

In cluster0,we can see more number of French restaurants,hotels,public places and so on.

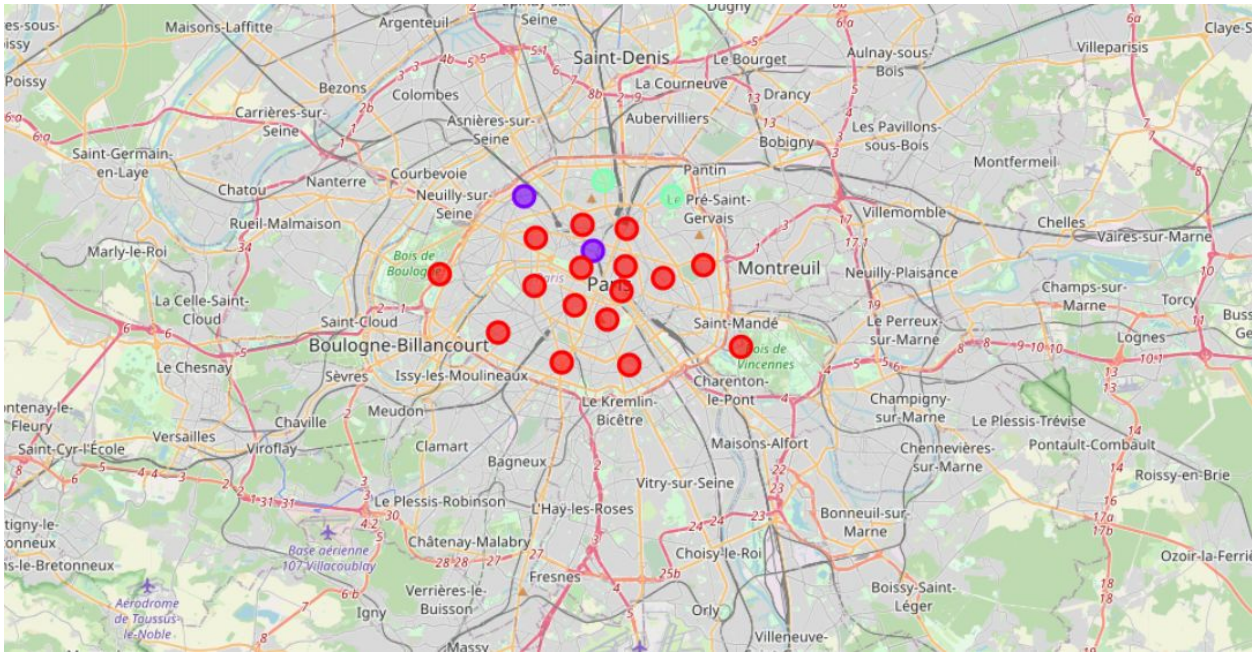
	Arrondissement_Num	Neighborhood	French_Name	Latitude	Longitude	Population_in_2019	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	Cluster Labels
7	2	Bourse	2eme Ardt	48.868279	2.342803	20260	French Restaurant	Wine Bar	Hotel	Cocktail Bar	Bakery	Coffee Shop	Bistro	Pedestrian Plaza	Cheese Shop	Ice Cream Shop	1
9	17	Batignolles-Monceau	17eme Ardt	48.887327	2.306777	167835	French Restaurant	Hotel	Italian Restaurant	Japanese Restaurant	Bakery	Bistro	Bus Stop	Restaurant	Plaza	Café	1

In cluster1,we can see French Restaurants,Wine bar,Hotels and so on.

	Arrondissement_Num	Neighborhood	French_Name	Latitude	Longitude	Population_in_2019	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	Cluster Labels
1	19	Buttes-Chaumont	19eme Ardt	48.887076	2.384821	186393	French Restaurant	Bar	Café	Seafood Restaurant	Beer Bar	Supermarket	Bistro	Hotel	Canal	Burger Joint	2
10	18	Buttes-Montmartre	18eme Ardt	48.892569	2.348161	195060	Bar	French Restaurant	Coffee Shop	Convenience Store	Café	Hotel	Restaurant	Sandwich Place	Cheese Shop	Pool	2

In cluster2,we can see bars,cafes,hotels and so on.

Now we can visualize the data using a folium map which represents the three clusters in three different colours.



Here we can see red color represents cluster0,violet color represents cluster1 and green color represents cluster2 .Each color represents each cluster label along with arrondissement names.Majority of the arrondissements fall under cluster0 while other arrondissements fall under cluster1 and cluster2.

Results and Discussion

Our data analysis and visualization using Folium maps and Foursquare data shown that a large number of Indian Restaurants are present in four arrondissements of Paris such as Entrepôt,Vaugirard,Buttes-Montmartre,and Menilmontant.Out of these four arrondissements,three arrondissements are in top five highly populated areas except for Entrepot.The number of Indian Restaurants are mostly present in this highly populated arrondissements.Considering the factor of competition we cannot setup a Indian restaurant in these areas.Our analysis on Farmer's Market of each arrondissement in Paris has shown that arrondissements Reuilly,Gobelins and Passy have most number of farmer

markets.For freshness and lower cost of ingredients,we can consider this locations for setting up a Indian Restaurant.A exploratory data analysis rendered nearby venues of each arrondissement in Paris.The analysis of each arrondissement provided the top most ten common venue categories of each arrondissement in determining the best suitable location for setting up the restaurant.The Clustering of arrondissements resulted in showing three clusters in which cluster0 shows most number of arrondissements with venues.As considering the factors of population,farmer's market,and most entertaining venues we can recommend the arrondissements Reuilly,Gobelins,and Passy for setting up the restaurant.But Gobelins has various number of restaurants and less entertaining public places,so we can recommend Reuilly or Passy for setting up a new Indian Restaurant as both of them have more entertaining public places.

Conclusion

This project helps in to have a better understanding of the neighborhoods with respect to the most common venues in that neighborhood.The analysis enabled us in figuring out two main regions for setting up the restaurant which are mostly visited by the tourists and locals.Further based analysis can provide which specific cuisine should be served by the restaurant for maximizing profits and it's up to the entrepreneur or stakeholder's final decision for choosing a recommended place,considering some features like cost of budget for setting up the restaurant.