Business density analysis in Paris-France

Obi Bertrand Obi
IBM Professional Certificate in Data Science
Cousera
obibertrand@gmail.com

Abstract

Choosing a small business location is an issue most entrepreneurs approach with a pinch of salt. Many small business startups, in order to reduce expenditure usually settle for a cheap location which may not attract enough clients to ensure its sustainability. In this project, we attempt to use business density analysis on neighborhoods in Paris in order to ensure optimal choice of location for small businesses. The results enable the determination of highly dense businesses in each neighborhood of Paris. Neighborhoods where particular businesses were not common could serve as favorable business location destinations after considering other factors. The project could also recommend favorable tourists destinations in the touristic city of Paris in France.

Introduction

1. Introduction

There are many things to consider when choosing a location for your business venture, whether setting up an office or a shop for the first time, or looking to expand into new areas. Following the recent release of a new interactive crime report which revealed fascinating insights about crime and security across the UK, entrepreneur Jake Fox reveals the key factors a business needs to consider when selecting a new location as; accessibility, security, competition, cost, availability of labor and growth potentials.

1.1. Problem statement

Even when Entrepreneurs possess the capital required in setting up a business, Chosing the right location that can attract potential customers for the business is a major decision to make. Even when a city is chosen, it requires further business density analysis or neighborhood analysis in order to choose a suitable neighborhood for set up of the business.

1.2. Project Objectives

The main objective of this project is to perform small business density analysis in the neighborhoods of the city of Paris in France as well as recommending favorable business locations for potential entrepreneurs.

Other objectives include the following:

- Build a data frame of neighborhoods in Paris - France by connecting to Open Data about Paris website
- Get the geographical coordinates of the neighborhoods in Paris using the geocoder function
- Obtain the venues data for the neighborhoods using Foursquare API
- Explore, cluster the neighborhoods and select the best cluster to open a new business
- Recommend favorable destinations for tourism in the city of Paris

1.3 Target audience

This project could be used by the following groups of people:

1. Business persons with interest in setting up small businesses in Paris
2. Customers looking for where to obtain particular services in Paris
3. Tourists interested in visiting and spending quality time in Paris for the first time
4. Paris city administration interested in channeling balance business set ups in particular neighborhoods in the city
5. Government interested in balance developments in the city of Paris.

Methodology

2.1. Background to study area

- Paris is the capital and most populous city of France, with an area of 105 square kilometres (41 square miles) and an official estimated population of 2,140,526 residents as of 1 January 2019. Since the 17th century, Paris is one of Europe's major centres of finance, diplomacy, commerce, fashion, science, and the arts. The geographical coordinates of Paris are: Latitute 48.8566101 and Logitude: 2.3514992. The City of Paris is the centre and seat of government of the Île-de-France, or Paris Region, which has an estimated official 2019 population of 12,213,364.
- This project will be implemented in Paris - France, owing to the fact that its the Capital city of France and a historic touristic city.

2.2. Data collection

- The following sources of data are used for the project:
- Csv file of open data on Paris-France showing neighborhoods, their population and geographical coordinates(Longitude and latitude)
- Use the Foursquare API to get all the business venues in the city center of Paris. There are 1817 venues divided in 223 unique categories
- Also use geocoder function to obtain geographical coordinates of all the venues obtained.

2.3. Data analyses

- The following data analysis procedures are used
- Explore the neighborhoods and segment the data using One hot encoding of the venues and categories of business
- Compute the top 5 most common business venues for each neighborhood(Arrondisement) in Paris
- Also compute the top 10 most common business venues in descending order for each neighborhood or arrondisement.
- Cluster the venues and represent on a Map using the Machine Learning KMeans clustering algorithm in scikit learn.

2.4. Project implementation tools

The Business density analysis project is implemented in Python 3.5, IBM Jupiter Lab using the Following Libraries:

- Numpy for Data analysis
- Pandas for Data Analysis
- Scikit learn for implementation of Machine learning Algorithms
- Matplotlib and seaborn for graphical representation of results.
- Folium for the production of Maps
- Geocoder from Geopy to convert an address into latitude and longitude values.
- Foursquare to obtain a data set of business venues in Paris

2.5. Presentation of results

- Table showing lists of neighborhoods(Arrondisements) and their geographical coordinates (latitude and longitudes)
- Map of Paris showing all neighborhoods
- Table showing business venues, neighborhoods and their geographical coordinates
- Lists of the top five most common business venues in each neighborhood and their density
- Table showing the top ten most common venues in each neighborhood(arrondisement)
- Line graph showing the elbow method of choosing the right number of clusters to segment the city of Paris according to business venues(k = 5)
- The map of Paris showing clusters of business venues and neighborhoods

Results

3.1 List of Neighborhoods in Paris and their geographical coordinates

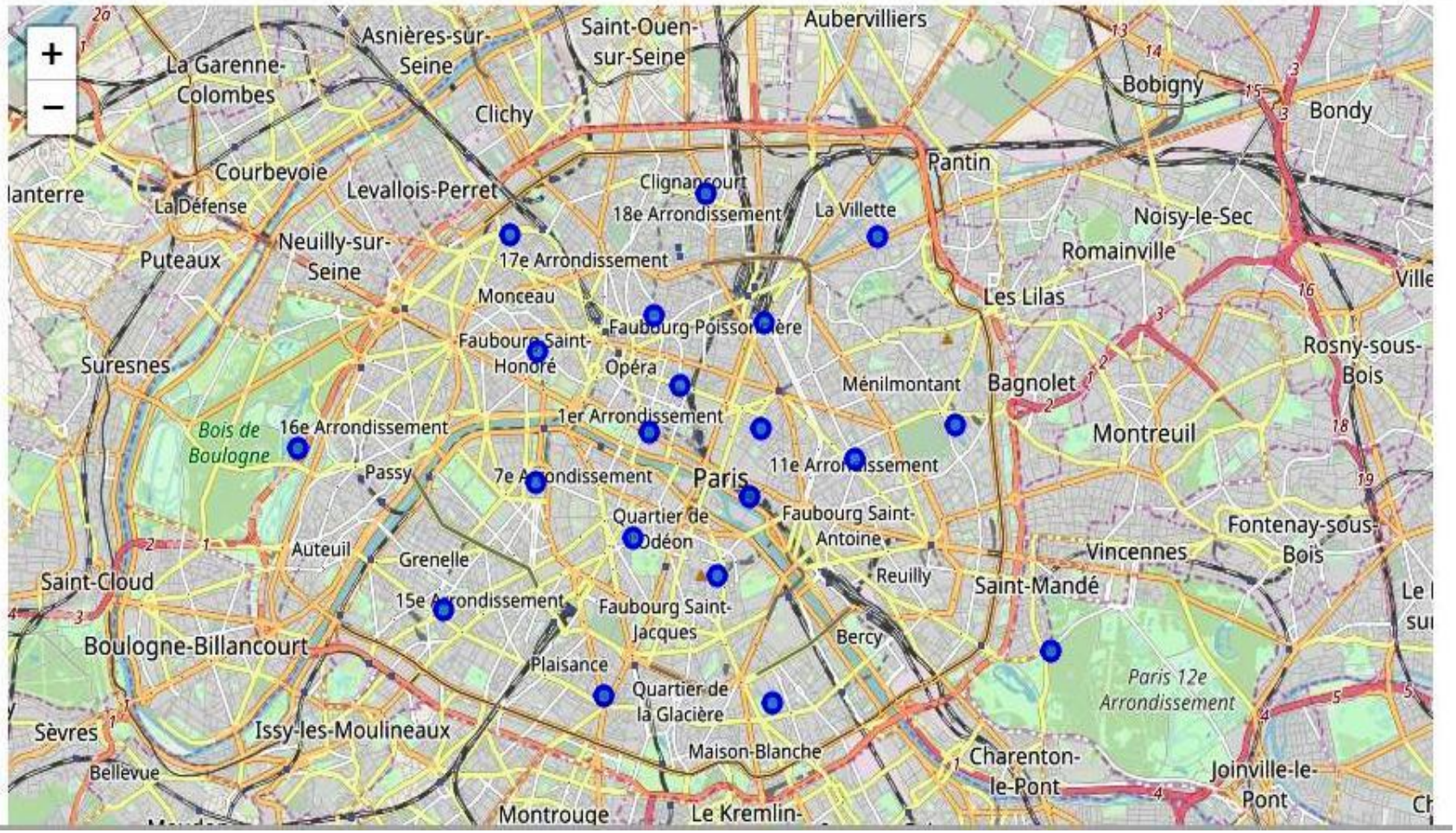
```
paris_data
```

[3]:

	Arrondissement	Neighborhood	Latitude	Longitude
0	1	Louvre	48.862563	2.336443
1	2	Bourse	48.868279	2.342803
2	3	Temple	48.862872	2.360001
3	4	Hôtel-de-Ville	48.854341	2.357630
4	5	Panthéon	48.844443	2.350715
5	6	Luxembourg	48.849130	2.332898
6	7	Palais-Bourbon	48.856174	2.312188
7	8	Elysée	48.872721	2.312554
8	9	Opéra	48.877164	2.337458
9	10	Entrepôt	48.876130	2.360728
10	11	Popincourt	48.859059	2.380058
11	12	Reuilly	48.834974	2.421325
12	13	Gobelins	48.828388	2.362272
13	14	Observatoire	48.829245	2.326542
14	15	Vaugirard	48.840085	2.292826

3.2 Map of Paris showing all Neighborhoods

[17]:



3.3 List of business venues and their geographical coordinates in Paris

[21]:
```
print(paris_venues.shape)
```

```
(1815, 8)
```

[22]:
```
print("There are {} unique categories.".format(len(paris_venues["Venue Category"].unique())))
```

```
There are 223 unique categories.
```

[23]:
```
paris_venues.head()
```

[23]:

	Arrondissement	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	1	Louvre	48.862563	2.336443	Musée du Louvre	48.860847	2.336440	Art Museum
1	1	Louvre	48.862563	2.336443	Comédie-Française	48.863088	2.336612	Theater
2	1	Louvre	48.862563	2.336443	Palais Royal	48.863758	2.337121	Historic Site
3	1	Louvre	48.862563	2.336443	Place du Palais Royal	48.862523	2.336688	Plaza
4	1	Louvre	48.862563	2.336443	La Clef Louvre Paris	48.863977	2.336140	Hotel

There are 1 815 Business venues made up of 223 unique categories in the 20 Neighborhoods (Arondisement) of Paris-France

3.4 Top five business venues in each neighborhood and their density

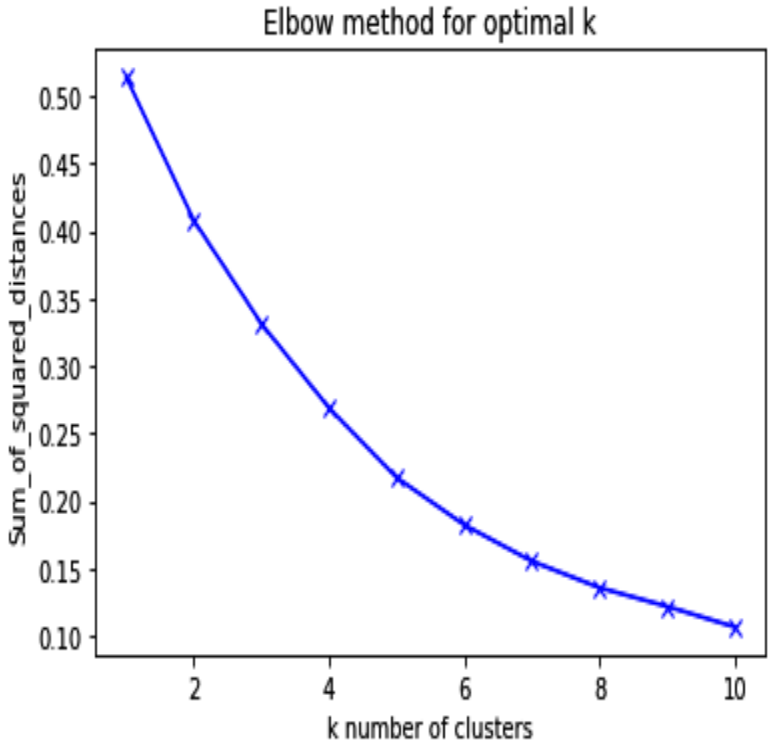
- Considering that Paris is a touristic city, Hotels, restaurants, Coffee shops, Bakery and Bars dominates most neighbourhoods.
- In Reuilly neigbourhood, despite enormous touristic potentials, restaurants and hotels seem to absent. This could be a great investment opportunity for an investor
- Banks also seem not to be very common in most communities this also represent a great investment opportunities.

3.5 Top ten most common venues in each Neighborhood (Arrondisement)

]:

ondissement	Neighborhood	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
17	Batignolles-Monceau	Italian Restaurant	French Restaurant	Hotel	Japanese Restaurant	Bakery	Bistro	Restaurant	Café	Plaza	Bar
2	Bourse	French Restaurant	Wine Bar	Bistro	Cocktail Bar	Bakery	Italian Restaurant	Japanese Restaurant	Pedestrian Plaza	Indie Movie Theater	Ice Cream Shop
19	Buttes-Chaumont	Bar	French Restaurant	Pizza Place	Hotel	Supermarket	Café	Smoke Shop	Restaurant	Bistro	Seafood Restaurant
18	Buttes-Montmartre	French Restaurant	Bar	Pizza Place	Italian Restaurant	Bistro	Café	Coffee Shop	Japanese Restaurant	Deli / Bodega	Park
8	Elysée	Hotel	French Restaurant	Italian Restaurant	Bakery	Art Gallery	Clothing Store	Japanese Restaurant	Theater	Bar	Thai Restaurant
10	Entrepôt	French Restaurant	Coffee Shop	Bistro	Pizza Place	Italian Restaurant	Bakery	Breakfast Spot	Indian Restaurant	Cocktail Bar	Seafood Restaurant
13	Gobelins	Vietnamese Restaurant	Thai Restaurant	Asian Restaurant	Chinese Restaurant	French Restaurant	Hotel	Bakery	Cantonese Restaurant	Supermarket	Japanese Restaurant
4	Hôtel-de-Ville	French Restaurant	Ice Cream Shop	Plaza	Wine Bar	Tapas Restaurant	Bakery	Pastry Shop	Cocktail Bar	Bookstore	Tea Room
1	Louvre	French Restaurant	Hotel	Plaza	Café	Japanese Restaurant	Exhibit	Historic Site	Art Museum	Cosmetics Shop	Udon Restaurant

3.6 Chosing the number of clusters to cluster the city of Paris according to neighborhoods



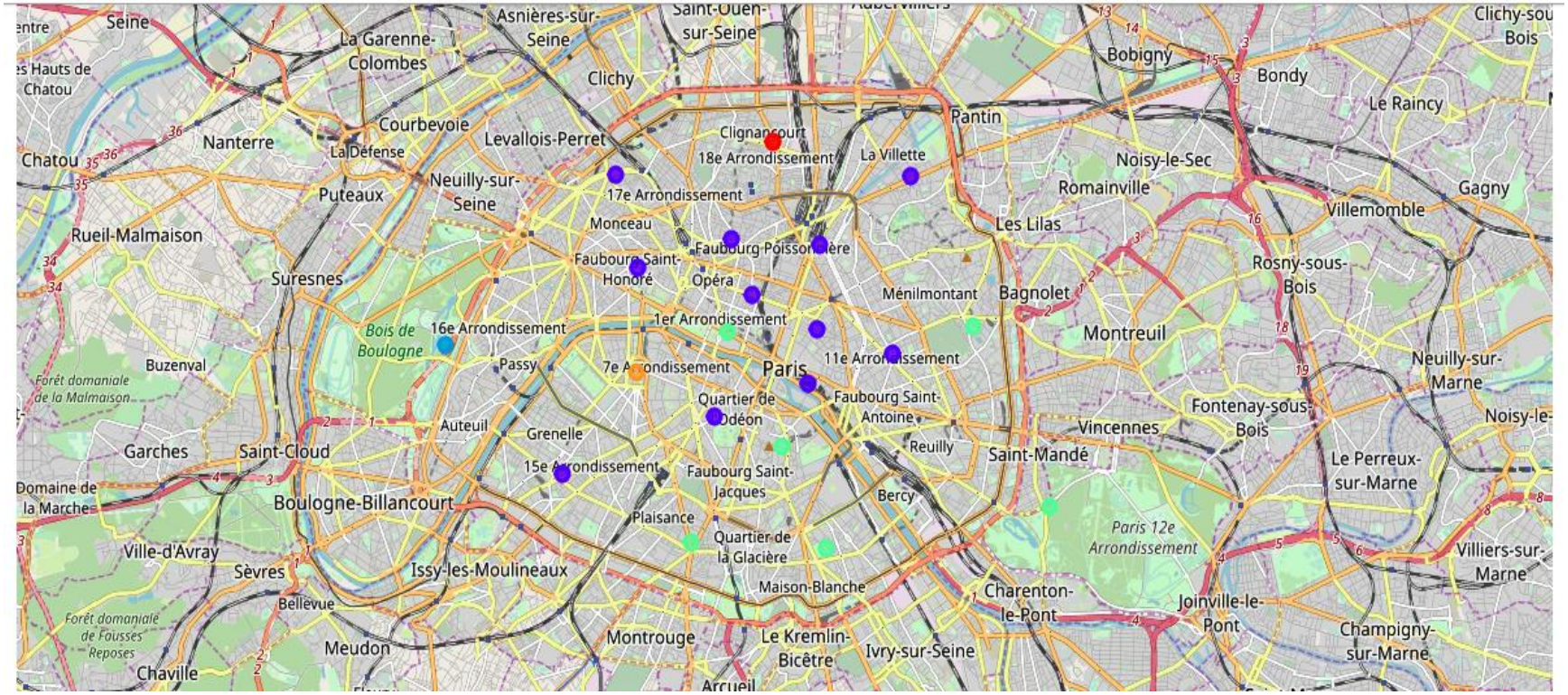
According to the figure above, the correct value for k could be 4 or 5. Here, we choose to cluster the data points using k=5.

3.7 Top ten most common venues in each Neighborhood (arrondisement) and their clusters

[19]:

	Arrondissement	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	10th Most Common Venue
0	1	Louvre	48.862563	2.336443	3	French Restaurant	Plaza	Hotel	Café	Japanese Restaurant	Exhibit	Historic Site	Udon Restaurant	Art Museum	Cosmetics Shop
1	2	Bourse	48.868279	2.342803	1	French Restaurant	Bistro	Wine Bar	Cocktail Bar	Italian Restaurant	Boutique	Japanese Restaurant	Bakery	Creperie	Hotel
2	3	Temple	48.862872	2.360001	1	French Restaurant	Art Gallery	Coffee Shop	Gourmet Shop	Wine Bar	Italian Restaurant	Chinese Restaurant	Moroccan Restaurant	Sandwich Place	Cocktail Bar
3	4	Hôtel-de-Ville	48.854341	2.357630	1	French Restaurant	Ice Cream Shop	Plaza	Wine Bar	Tapas Restaurant	Pastry Shop	Bakery	Cocktail Bar	Clothing Store	Gastropub
4	5	Panthéon	48.844443	2.350715	3	French Restaurant	Bar	Wine Bar	Café	Hotel	Bakery	Plaza	Coffee Shop	Greek Restaurant	Museum
5	6	Luxembourg	48.849130	2.332898	1	French Restaurant	Hotel	Italian Restaurant	Wine Bar	Chocolate Shop	Bistro	Seafood Restaurant	Tea Room	Plaza	Pastry Shop
6	7	Palais-Bourbon	48.856174	2.312188	4	French Restaurant	Hotel	Plaza	Italian Restaurant	Café	Cocktail Bar	History Museum	Historic Site	Ice Cream Shop	Bar
7	8	Elysée	48.872721	2.312554	1	French Restaurant	Hotel	Art Gallery	Italian Restaurant	Bakery	Theater	Clothing Store	Japanese Restaurant	Thai Restaurant	Boutique
8	9	Opéra	48.877164	2.337458	1	French Restaurant	Hotel	Cocktail Bar	Italian Restaurant	Bistro	Bar	Bakery	Wine Bar	Lounge	Pizza Place
9	10	Entrepôt	48.876130	2.360728	1	French	Coffee	Bistro	Pizza Place	Japanese	Italian	Bakery	Breakfast	Indian	Cocktail Bar

3.8 The map of Paris showing clusters of business venues and neighborhoods



There are five clusters with five different colors (Cluster1: Red, Cluster2: Purple, Cluster3: Blue, Cluster4: Green, Cluster5: Orange) on the map. The Jupiter notebook presents a better map that can be zoomed to view all the clusters perfectly.

Clusters and Recommendations

Cluster 1: Red

```
[21]: paris_merged.loc[paris_merged["Cluster Labels"] == 0, paris_merged.columns[[1] + list(range(5, paris_merged.shape[1]))]]
```

[21]:

	Neighborhood	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
17	Buttes-Montmartre	French Restaurant	Bar	Pizza Place	Bistro	Italian Restaurant	Coffee Shop	Café	Art Gallery	Middle Eastern Restaurant	Convenience Store

Most common venues in Cluster 1 are restauration services consisting of restaurants, Bars and Café shops. This represents 80% of the venues. Opportunities exists here for the establishment of other businesses like Hotels, banks, book stores, education services, Insurance, Shopping Centers, etc.

Cluster 2: Purple

[23]:
```
paris_merged.loc[paris_merged["Cluster Labels"] == 1, paris_merged.columns[[1] + list(range(5, paris_merged.shape[1]))]]
```

[23]:

	Neighborhood	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
1	Bourse	French Restaurant	Bistro	Wine Bar	Cocktail Bar	Italian Restaurant	Boutique	Japanese Restaurant	Bakery	Creperie	Hotel
2	Temple	French Restaurant	Art Gallery	Coffee Shop	Gourmet Shop	Wine Bar	Italian Restaurant	Chinese Restaurant	Moroccan Restaurant	Sandwich Place	Cocktail Bar
3	Hôtel-de-Ville	French Restaurant	Ice Cream Shop	Plaza	Wine Bar	Tapas Restaurant	Pastry Shop	Bakery	Cocktail Bar	Clothing Store	Gastropub
5	Luxembourg	French Restaurant	Hotel	Italian Restaurant	Wine Bar	Chocolate Shop	Bistro	Seafood Restaurant	Tea Room	Plaza	Pastry Shop
7	Elysée	French Restaurant	Hotel	Art Gallery	Italian Restaurant	Bakery	Theater	Clothing Store	Japanese Restaurant	Thai Restaurant	Boutique
8	Opéra	French Restaurant	Hotel	Cocktail Bar	Italian Restaurant	Bistro	Bar	Bakery	Wine Bar	Lounge	Pizza Place
9	Entrepôt	French Restaurant	Coffee Shop	Bistro	Pizza Place	Japanese Restaurant	Italian Restaurant	Bakery	Breakfast Spot	Indian Restaurant	Cocktail Bar
10	Popincourt	Bar	French Restaurant	Cocktail Bar	Bistro	Pizza Place	Italian Restaurant	Restaurant	Vegetarian / Vegan Restaurant	Beer Bar	Wine Bar

Cluster 2 has eleven neighbourhoods consisting of mainly restauration services and hotels. This could also be a great location for the establishment of shopping centers, bakery and banks

Cluster 3: Blue

[24]:
```
paris_merged.loc[paris_merged["Cluster Labels"] == 2, paris_merged.columns[[1] + list(range(5, paris_merged.shape[1]))]]
```

[24]:

	Neighborhood	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
15	Passy	Plaza	Lake	Pool	French Restaurant	Garden	Diner	Skate Park	Bus Station	Bus Stop	Cafeteria

Cluster 3 also has one neighbourhood consisting of mainly restauration services and touristic sites. Business like banks or payment counters and hotels could also be established here

Cluster 3 also has one neighbourhood consisting of mainly restauration services and touristic sites. Business like banks or payment counters and hotels could also be established here

Cluster 4: Green

[25]:
```
paris_merged.loc[paris_merged["Cluster Labels"] == 3, paris_merged.columns[[1] + list(range(5, paris_merged.shape[1]))]]
```

[25]:

	Neighborhood	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	Louvre	French Restaurant	Plaza	Hotel	Café	Japanese Restaurant	Exhibit	Historic Site	Udon Restaurant	Art Museum	Cosmetics Shop
4	Panthéon	French Restaurant	Bar	Wine Bar	Café	Hotel	Bakery	Plaza	Coffee Shop	Greek Restaurant	Museum
11	Reuilly	Lake	Zoo	Diner	Bus Stop	Exhibit	French Restaurant	Monument / Landmark	Japanese Restaurant	Hot Dog Joint	Fast Food Restaurant
12	Gobelins	Vietnamese Restaurant	Thai Restaurant	Asian Restaurant	French Restaurant	Chinese Restaurant	Bakery	Hotel	Cantonese Restaurant	Japanese Restaurant	Cambodian Restaurant
13	Observatoire	French Restaurant	Hotel	Italian Restaurant	Bistro	Café	Bar	Vietnamese Restaurant	Plaza	Bakery	Pizza Place
19	Ménilmontant	French Restaurant	Bar	Bakery	Plaza	Bistro	Italian Restaurant	Café	Park	Bookstore	Supermarket

Cluster 4 has six Neighborhoods consisting of mainly restauration services, touristic sites and few hotels. Hotels could be a great investment in this cluster. Other posiible investments are shopping centers, retails shops, banks, etc.

Cluster 5: Orange (west of Paris)

```
[26]: paris_merged.loc[paris_merged["Cluster Labels"] == 4, paris_merged.columns[[1] + list(range(5, paris_merged.shape[1]))]]
```

[26]:

	Neighborhood	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
6	Palais-Bourbon	French Restaurant	Hotel	Plaza	Italian Restaurant	Café	Cocktail Bar	History Museum	Historic Site	Ice Cream Shop	Bar

Cluster 5 has one Neighborhood mainly consisting of restauration, Hotels and touristic business venues. Additional restaurants, banks and retail shops could be established here.

Cluster 5 has one Neighborhood mainly consisting of restauration, Hotels and touristic business venues. Additional restaurants, banks and retail shops could be established here.

4. Conclusion

Using Business density analysis, lots of opportunities exists to established competitive businesses in various neighborhoods of the city of Paris. This analysis can also be conducted in other cities in the World with adequate venues data in Foursquare. The results of this analysis when combined with other factors will enable improved business locations for sustainability.