The Battle of the Neighborhoods - Report

1 Introduction & Business Problem

The city of Paris is the most populated agglomeration in France and one if not the most attractive in Europe. Paris is a lively, diverse, dynamic city and is the financial as well as political capital of France.
It is a multicultural city which provides a lot of opportunities as well as a favorable business environment.

The city has attracted many players in the market. It is a global hub of business and commerce. The city is a major center for banking and financial (one of the most important in Europe), retail trade, world trade, transport, tourism, real estate, new media, traditional media, advertising, legal services, accounting, insurance, theater, fashion and arts in Europe. It also means that the market is very competitive. As it is a very developed city, the cost of business is also one of the highest.

Thus, any new commercial enterprise or expansion must be carefully analyzed. The information drawn from the analysis will allow a good understanding the business environment that helps strategically target the market. This analysis will help in risk reduction. And the return on investment will be reasonable, though.

1.1 Political Division of Paris

The City of Paris like any big city is divided in many, many ways. The city itself was divided into 20 different Faubourgs and later into **Arrondissments** (districts) for the purpose of taxation and policing, but there were social divisions as well that were just as concrete. These divisions were influenced by the differing ideals between social groups. Understanding these divisions is important to both the political and social history of Paris.

You will find in this report a map showing Paris arrondissements and a map presenting Paris neighborhoods. Paris arrondissements and districts are the administrative divisions that share Paris in 20 municipals arrondissements. Within each of its boroughs, the neighborhoods are making the charm and the reputation of the city of Paris.

The twenty arrondissements of Paris all have their own character through their tourist sites, places to go out, restaurants, and homes. Each Parisian arrondissement has four administrative districts.

Each of Paris' 20 administrative districts (or arrondissements) are officially divided into 4 quartiers.

1.2 Business Problem

City of Paris is famous for its excellent cuisine. Its food culture includes an array of international cuisines influenced by the city's immigrant history. Chinese restaurants have become so popular in France now it seems that there is one on every corner, not only in major cities but also in smaller cities.

Starting a Chinese restaurant can be a great business opportunity, but you need to distinguish yourself from others to enjoy long-term success.

If you plan a real restaurant that can demand higher prices for fresh fish or fresh foods, delivered daily from Japan or China, focus on neighborhoods and outlets that already attract a sophisticated Chinese or others client. If you plan a cheap buffet restaurant, points to the masses looking for affordable high-traffic locations with large shopping centers and other local points of interest.

My client wants to open their business in Paris area, not necessary in the center of Paris. So according to his need, I focus on that borough during my analysis. We define potential neighborhood based on the number of Chinese bars which are operating right in each neighborhood. Paris has full potential but also is a very challenging district to open a business because of high competition. New Chinese bar should be open in an area that inadequate neighborhood in this way the bar can attract more customers. Therefore, this analysis necessary to ensure that we have enough customers and that we are not so close to other Chinese places, and if we are close to others, to provide the best food in order to attract as many customers as possible.

2 Data acquisition and cleaning

As specify earlier, Paris has a total of 20 boroughs and 80 neighborhoods. In order to segment the neighborhoods and explore them, we will essentially need a dataset that contains the 20 boroughs and the neighborhoods that exist in each borough as well as the latitude and longitude coordinates of each neighborhood.

The first dataset consists of the 20 boroughs with their respective geographical coordinates and are available at:

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

The second dataset consists of the 80 neighborhoods with their respective geographical coordinate, the boroughs to witch the belong and are available at:

- https://opendata.paris.fr/explore/embed/dataset/quartier_paris/table/

Merging those two datasets, we finally have our dataset with all 20 boroughs and 4 neighborhoods in each.

	Borough	Neighborhood	Latitude	Longitude
0	Bourse	Mail	48.8680083374	2.34469912743
1	Bourse	Bonne-Nouvelle	48.8671501183	2.35008019041
2	Bourse	Gaillon	48.8693066381	2.33343180766
3	Bourse	Vivienne	48.8691001998	2.33946074375
4	Gobelins	Gare	48.8275274578	2.37239773692

Figure 1 - Boroughs, Neighborhoods of Paris

Why using geographical data?

- Paris city geographical coordinates data will be utilized as input for the *Foursquare API*[1], that will be leveraged to provision venues information for each neighborhood.

- We will use the Foursquare API to explore neighborhoods in Paris.

In addition, Chinese category Id **4bf58dd8d48988d145941735**[2] is used for retrieving data from Foursquare API.

3 Methodology

In this project, I will use the basic methodology. Given the dataset with Boroughs, Neighborhoods and their respective geographical coordinates, we will use the *Foursquare API* to explore neighborhoods of Paris. After that, explore function to get Chinese restaurant categories in each neighborhood. Finally, I will use K-Means[3] clustering to cluster and find the best places(venues) to build a Chinese restaurant in Paris.

For the sake of simplicity, we will limit our research to boroughs like **Louvre, Bourse, Temple, Hôtel-de-Ville, Panthéon, Luxembourg and Palais-Bourbon.**

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue distance	Venue Category
0	Mail	48.8680083374	2.34469912743	Mr Zhao	48.869766	2.343644	210	Chinese Restaurant
1	Mail	48.8680083374	2.34469912743	Au Piment du Centre	48.859342	2.348966	1014	Chinese Restaurant
2	Mail	48.8680083374	2.34469912743	Hao Long	48.865660	2.336954	624	Chinese Restaurant
3	Mail	48.8680083374	2.34469912743	Traiteur Chez Zhang	48.861904	2.344043	681	Chinese Restaurant
4	Mail	48.8680083374	2.34469912743	Tafa	48.867023	2.336098	639	Chinese Restaurant

```
paris_venues_cn.shape
```

```
(1228, 8)
```

Results returns by *Foursquare API* give us information's about neighborhood (geographical coordinates), venue (geographical coordinates and different category).

- The most important thing here is the **hierarchical list of categories applied to venues.**

For our research perimeter, we find more than 19 unique categories.

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

```
There are 19 unique categories.
```

We then use those 19 unique categories to One-Hot[4] encode each Venue category values.

---

[1] The **Foursquare API** provides location-based experiences with diverse information about venues, users, photos, and check-ins.

[2] **Foursquare API** categories list are available at https://developer.foursquare.com/docs/resources/categories

[3] https://www.wikiwand.com/en/K-means_clustering

[4] https://www.wikiwand.com/en/One-hot

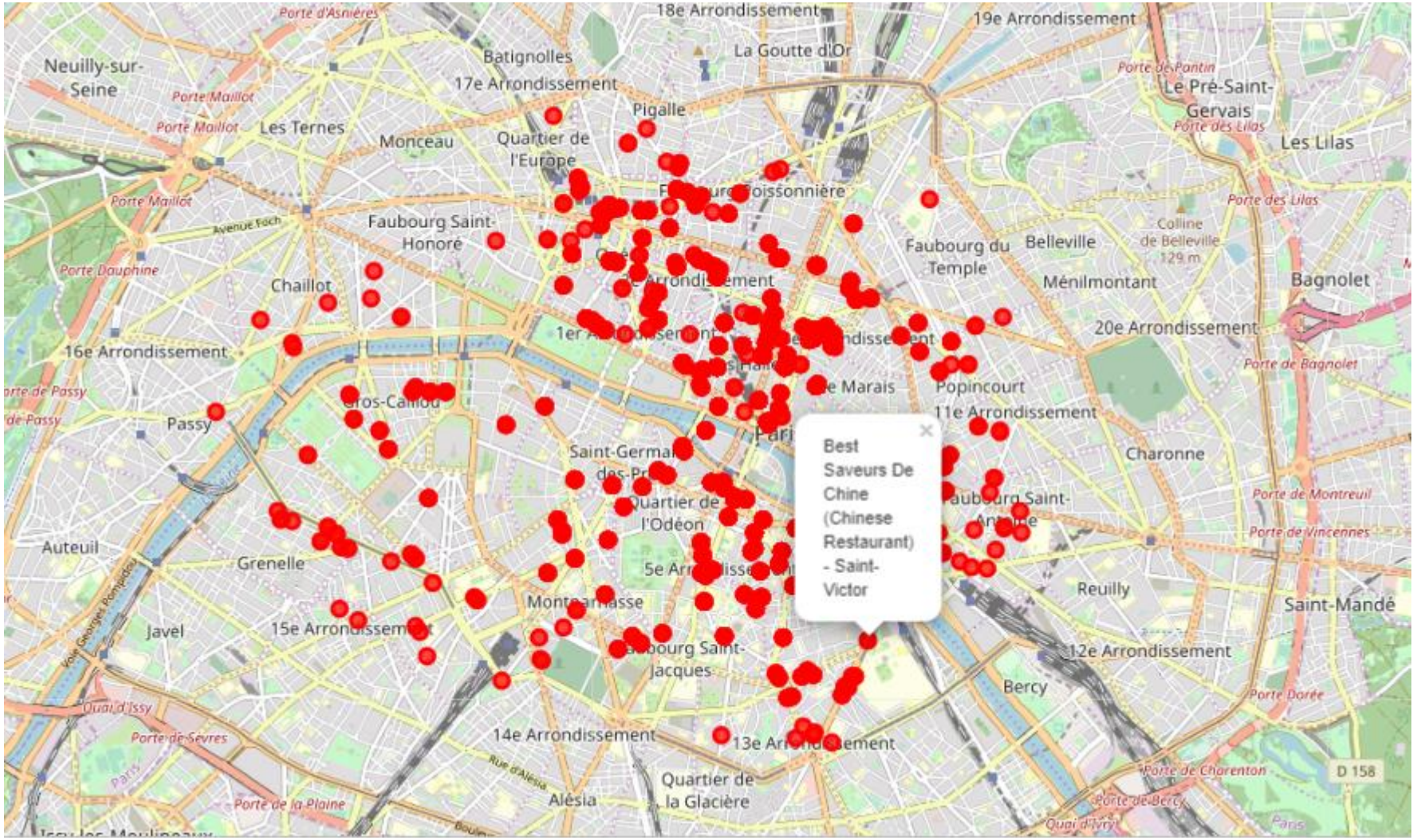

Figure 2 - Chinese restaurants in Paris (Louvre, Bourse, Temple, Hôtel-de-Ville, Panthéon, Luxembourg and Palais-Bourbon).

	Neighborhood	Asian Restaurant	BBQ Joint	Bubble Tea Shop	Cantonese Restaurant	Chinese Restaurant	Dessert Shop	Dim Sum Restaurant	Dongbei Restaurant	Fondue Restaurant	Karaoke Bar	Noodle House	Shandong Restaurant	Shanxi Restaurant	Sushi Restaurant
0	Mail	0	0	0	0	1	0	0	0	0	0	0	0	0	0
1	Mail	0	0	0	0	1	0	0	0	0	0	0	0	0	0
2	Mail	0	0	0	0	1	0	0	0	0	0	0	0	0	0
3	Mail	0	0	0	0	1	0	0	0	0	0	0	0	0	0
4	Mail	0	0	0	0	1	0	0	0	0	0	0	0	0	0

Figure 3 -One-Hot Encoding Venue Categories by Neighborhood.

We then use this feature to group the neighborhoods into clusters. *K-means* clustering algorithm with 5 clusters will be used to complete this task. *Folium* library will also be used to visualize the neighborhoods in Paris and its emerging clusters.

	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
0	Archives	Chinese Restaurant	Szechuan Restaurant	Asian Restaurant	Noodle House	Dim Sum Restaurant	BBQ Joint	Cantonese Restaurant
1	Arsenal	Chinese Restaurant	Szechuan Restaurant	Asian Restaurant	BBQ Joint	Cantonese Restaurant	Noodle House	Dim Sum Restaurant
2	Arts-et-Métiers	Chinese Restaurant	Szechuan Restaurant	Asian Restaurant	Noodle House	Dim Sum Restaurant	Shanxi Restaurant	Dongbei Restaurant
3	Bonne-Nouvelle	Chinese Restaurant	Asian Restaurant	Noodle House	Dim Sum Restaurant	Taiwanese Restaurant	Szechuan Restaurant	Tea Room
4	Ecole-Militaire	Chinese Restaurant	Asian Restaurant	Dessert Shop	Fondue Restaurant	BBQ Joint	Bubble Tea Shop	Cantonese Restaurant

Figure 4 - 7 of 10 Most Common Venue

Applying *K-means*, each borough, neighborhood as well as their respective Latitude and Longitude will be grouped in each cluster. A plot will be made to visualize all those results.

	Borough	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
0	Bourse	Mail	48.8680083374	2.34469912743	2	Chinese Restaurant	Dim Sum Restaurant	Noodle House	Asian Restaurant	Szechuan Restaurant	Tea Room	Taiwanese Restaurant
1	Bourse	Bonne-Nouvelle	48.8671501183	2.35008019041	2	Chinese Restaurant	Asian Restaurant	Noodle House	Dim Sum Restaurant	Taiwanese Restaurant	Szechuan Restaurant	Tea Room
2	Bourse	Gaillon	48.8693066381	2.33343180766	2	Chinese Restaurant	Noodle House	Dim Sum Restaurant	Szechuan Restaurant	Asian Restaurant	Taiwanese Restaurant	Tea Room
3	Bourse	Vivienne	48.8691001998	2.33946074375	2	Chinese Restaurant	Asian Restaurant	Dim Sum Restaurant	Szechuan Restaurant	Noodle House	Taiwanese Restaurant	Tea Room
4	Hôtel-de-Ville	Arsenal	48.851585175	2.36476795387	4	Chinese Restaurant	Szechuan Restaurant	Asian Restaurant	BBQ Joint	Cantonese Restaurant	Noodle House	Dim Sum Restaurant
5	Hôtel-de-Ville	Saint-Merri	48.8585213723	2.35166696714	4	Chinese Restaurant	Noodle House	Asian Restaurant	Szechuan Restaurant	Tea Room	Cantonese Restaurant	Dim Sum Restaurant

Figure 5- Borough, Neighborhood grouped in clusters.

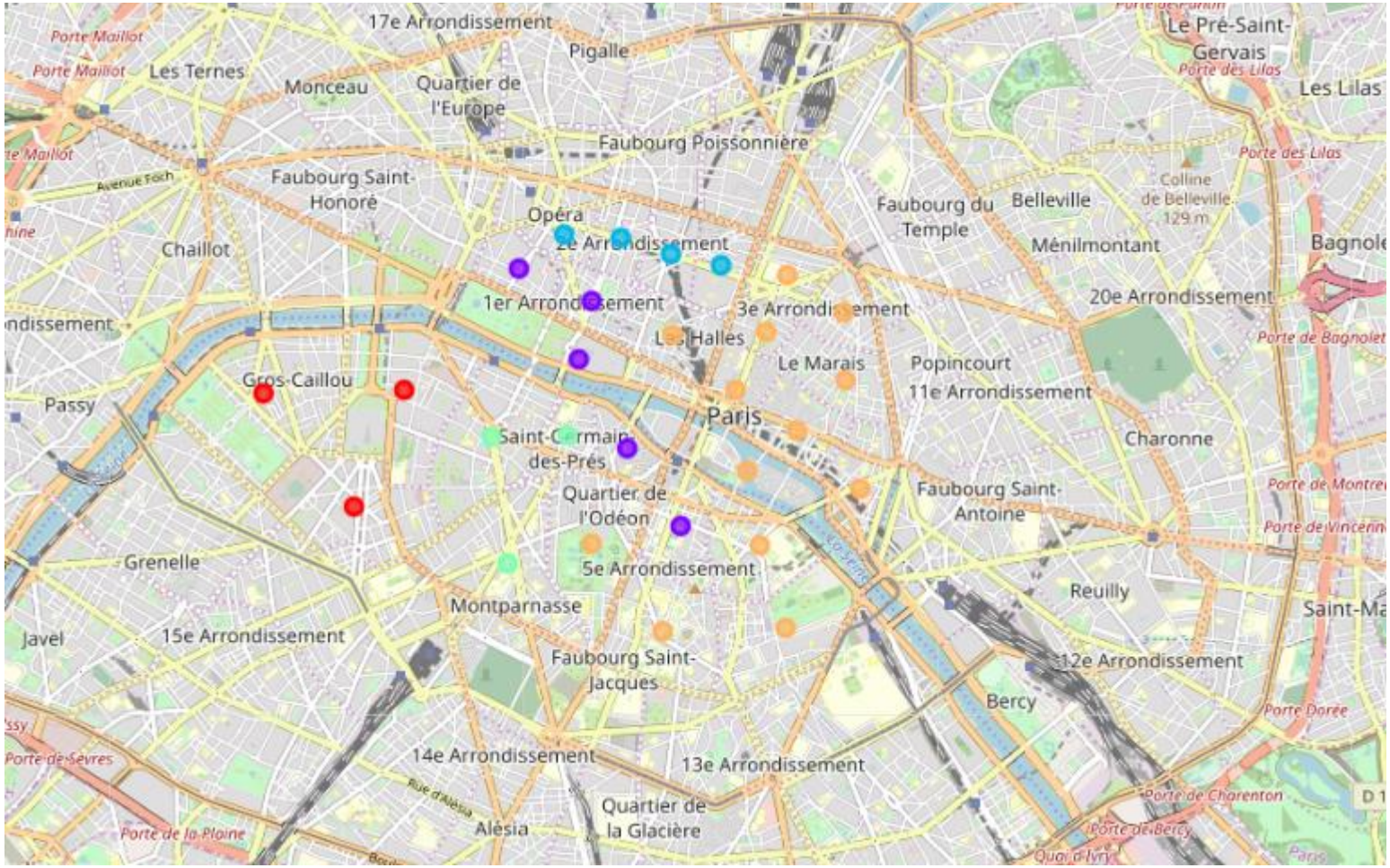

Figure 6 - Results with 5 clusters

4 Analyze results

K-means cluster our data into 5 clusters and return those items.

4.1 Cluster 0

	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
12	Invalides	Chinese Restaurant	Asian Restaurant	Dim Sum Restaurant	Fondue Restaurant	BBQ Joint	Bubble Tea Shop	Cantonese Restaurant
13	Ecole-Militaire	Chinese Restaurant	Asian Restaurant	Dessert Shop	Fondue Restaurant	BBQ Joint	Bubble Tea Shop	Cantonese Restaurant
14	Gros-Caillou	Chinese Restaurant	Asian Restaurant	Dim Sum Restaurant	Fondue Restaurant	BBQ Joint	Bubble Tea Shop	Cantonese Restaurant

4.2 Cluster 1

	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
9	Sorbonne	Chinese Restaurant	Asian Restaurant	Szechuan Restaurant	Dim Sum Restaurant	Noodle House	Cantonese Restaurant	Karaoke Bar
21	Palais-Royal	Chinese Restaurant	Tea Room	Taiwanese Restaurant	Noodle House	Asian Restaurant	Dim Sum Restaurant	Szechuan Restaurant
22	Saint-Germain-l'Auxerrois	Chinese Restaurant	Dim Sum Restaurant	Taiwanese Restaurant	Asian Restaurant	Tea Room	Szechuan Restaurant	Sushi Restaurant
23	Place-Vendôme	Chinese Restaurant	Szechuan Restaurant	Taiwanese Restaurant	Noodle House	Vietnamese Restaurant	Dim Sum Restaurant	Sushi Restaurant
25	Monnaie	Chinese Restaurant	Noodle House	Dim Sum Restaurant	Tea Room	Szechuan Restaurant	Cantonese Restaurant	Vietnamese Restaurant

4.3 Cluster 2

	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
0	Mail	Chinese Restaurant	Dim Sum Restaurant	Noodle House	Asian Restaurant	Szechuan Restaurant	Tea Room	Taiwanese Restaurant
1	Bonne-Nouvelle	Chinese Restaurant	Asian Restaurant	Noodle House	Dim Sum Restaurant	Taiwanese Restaurant	Szechuan Restaurant	Tea Room
2	Gaillon	Chinese Restaurant	Noodle House	Dim Sum Restaurant	Szechuan Restaurant	Asian Restaurant	Taiwanese Restaurant	Tea Room
3	Vivienne	Chinese Restaurant	Asian Restaurant	Dim Sum Restaurant	Szechuan Restaurant	Noodle House	Taiwanese Restaurant	Tea Room

4.4 Cluster 3

	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
15	Saint-Thomas-d'Aquin	Chinese Restaurant	Asian Restaurant	Fondue Restaurant	Dim Sum Restaurant	BBQ Joint	Bubble Tea Shop	Cantonese Restaurant
24	Saint-Germain-des-Prés	Chinese Restaurant	Asian Restaurant	Cantonese Restaurant	Dim Sum Restaurant	Noodle House	Fondue Restaurant	Karaoke Bar
27	Notre-Dame-des-Champs	Chinese Restaurant	Asian Restaurant	Dim Sum Restaurant	Fondue Restaurant	Cantonese Restaurant	Noodle House	BBQ Joint

4.5 Cluster 4

	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
4	Arsenal	Chinese Restaurant	Szechuan Restaurant	Asian Restaurant	BBQ Joint	Cantonese Restaurant	Noodle House	Dim Sum Restaurant
5	Saint-Merri	Chinese Restaurant	Noodle House	Asian Restaurant	Szechuan Restaurant	Tea Room	Cantonese Restaurant	Dim Sum Restaurant
6	Notre-Dame	Chinese Restaurant	Szechuan Restaurant	Asian Restaurant	Cantonese Restaurant	Noodle House	Tea Room	Dim Sum Restaurant
7	Saint-Gervais	Chinese Restaurant	Szechuan Restaurant	Asian Restaurant	Cantonese Restaurant	BBQ Joint	Noodle House	Dim Sum Restaurant
8	Jardin-des-Plantes	Chinese Restaurant	Szechuan Restaurant	Shandong Restaurant	Noodle House	Dongbei Restaurant	Bubble Tea Shop	Dim Sum Restaurant
10	Val-de-Grâce	Chinese Restaurant	Noodle House	Szechuan Restaurant	Shandong Restaurant	Dim Sum Restaurant	Vietnamese Restaurant	Dongbei Restaurant
11	Saint-Victor	Chinese Restaurant	Asian Restaurant	Szechuan Restaurant	Noodle House	Dim Sum Restaurant	Cantonese Restaurant	Shandong Restaurant
16	Sainte-Avoie	Chinese Restaurant	Szechuan Restaurant	Asian Restaurant	Noodle House	Dim Sum Restaurant	Tea Room	Dongbei Restaurant
17	Arts-et-Métiers	Chinese Restaurant	Szechuan Restaurant	Asian Restaurant	Noodle House	Dim Sum Restaurant	Shanxi Restaurant	Dongbei Restaurant
18	Archives	Chinese Restaurant	Szechuan Restaurant	Asian Restaurant	Noodle House	Dim Sum Restaurant	BBQ Joint	Cantonese Restaurant
19	Enfants-Rouges	Chinese Restaurant	Szechuan Restaurant	Asian Restaurant	Noodle House	Dim Sum Restaurant	Cantonese Restaurant	Shanxi Restaurant
20	Halles	Chinese Restaurant	Szechuan Restaurant	Asian Restaurant	Dim Sum Restaurant	Taiwanese Restaurant	Noodle House	Tea Room

Based on dataframe analysis above

- **Cluster 0 (Invalides, Ecole-Militaire, Gros-Caillou**) and
- **Cluster 3 (Saint-Thomas d'Acquin, Saint-Germain-des-Près, Notre-Dame-Des-Champs)**

areas are the best places to open a new Chinese restaurant business.

5 Discussion

In this section, I will focus on what we have noticed and the recommendation that I can make based on previous results.

- This analysis is performed on one hand on limited data and on other on reduced perimeter. We have focused our research around Louvre, Bourse, Temple, Hôtel-de-Ville, Panthéon, Luxembourg and Palais-Bourbon. This may be right or not, but if good amount of data were available there is scope to come up with better results.

- There is high competition around Notre-Dame, Jardins-des-Plantes and Val-de-Grace so it is very risky to open business in these areas.

- Sorbonne has also potential where it's closes to University and high schools.

- We can more detailed analysis by adding some factors such as transportation, demographics of inhabitants.

Finally, *Foursquare API* proved to be a good source of data but frustrating at times. Despite having a Developer account, I regularly exceeded my hourly limit locking me out for the day.

6 Conclusion

Although the goals of this project were met there is a lot to do to improve it.

In this study, I analyzed the relationship between business development and some useful machine learning technics such as clustering to group a set of Chinese restaurants in such a way that those restaurants in the same group (called a **cluster**) are more similar (in some sense) to each other than to those in other groups (clusters).

I also performed data analysis through this information by adding geographical coordinates of districts and some Chinese restaurants as static data on GitHub. In future studies, these data can also be accessed dynamically from specific platforms or packages.

I ended the study by visualizing the data and clustering information on the Paris map. In future studies, web or telephone applications can be carried out to direct investors.

With these analyses and information's, my client can choose the place where open his Chinese restaurant.

7 References

- Forsquare API https://developer.foursquare.com/
- Coursera https://www.coursera.org/professional-certificates/ibm-data-science
- Paris https://www.wikiwand.com/en/Paris
- Arrondissements of Paris https://www.wikiwand.com/en/Arrondissements_of_Paris
- Paris Neighborhoods https://www.parisinsidersguide.com/paris-neighborhoods.html