

Capstone project

Battle of neighborhoods

Opening of a new restaurant in
Paris



Introduction

Paris is the capital and most populous city of France. Since the 17th century, Paris has been one of Europe's major centres of finance, diplomacy, commerce, fashion, science and arts.

Project :

For the Capstone project, I chose the city of Paris to analyze its geospatial data in order to select a location for a new restaurant.

Criteria :

The best places to open a new Japanese restaurant are areas that are near places where people wait to eat : Art gallery, spa, clothing stores. The client's criteria is to have 2 out of three of these 3 places.

In addition to that, a place where there is a very low concurrence.



Battle of neighborhoods

Data cleaning

Paris is divided into 20 Municipal Districts (or administrative districts), reduced to arrondissements. They are normally referred to by the district number rather than by a name.

The data used is available on the web and can be manipulated and cleaned to provide a meaningful dataset to use. the data used is from : <https://www.data.gouv.fr/fr/datasets/arrondissements-1/>

The data have been cleaned using BeautifulSoup.
On the right you can see its shape

	Arrondissement Num	french name	Neighborhood	lat	lon
0	1	1er Ardt	Louvre	48.8626627018	2.33644336205
1	2	2ème Ardt	Bourse	48.8682792225	2.34280254689
2	17	17ème Ardt	Batignolles-Monceau	48.887326522	2.30877899057
3	14	14ème Ardt	Observatoire	48.8282445005	2.3265420442
4	20	20ème Ardt	Ménilmontant	48.8634605789	2.40118812928
5	7	7ème Ardt	Palais-Bourbon	48.8561744286	2.31218769146
6	11	11ème Ardt	Popincourt	48.8590592213	2.3800583082
7	13	13ème Ardt	Gobelins	48.8288880317	2.36227244042
8	4	4ème Ardt	Hôtel-de-Ville	48.8543414263	2.35762962032
9	8	8ème Ardt	Élysée	48.8727206374	2.3125540224
10	18	18ème Ardt	Buttes-Montmartre	48.892569268	2.34816051956
11	9	9ème Ardt	Opéra	48.8771835173	2.33745754348
12	19	19ème Ardt	Buttes-Chaumont	48.8870759966	2.38482096015
13	15	15ème Ardt	Vaugirard	48.8400353759	2.29282582242
14	3	3ème Ardt	Temple	48.86287236	2.3600009859
15	5	5ème Ardt	Panthéon	48.8444431505	2.35071480958
16	6	6ème Ardt	Luxembourg	48.8491303586	2.33289799905
17	12	12ème Ardt	Reully	48.8349743315	2.42132490078
18	10	10ème Ardt	Entrepôt	48.8761300365	2.36072948785
19	16	16ème Ardt	Passy	48.8603921054	2.26197078836

Data Vizualisation

Different neighborhoods in paris are Visualized as below:



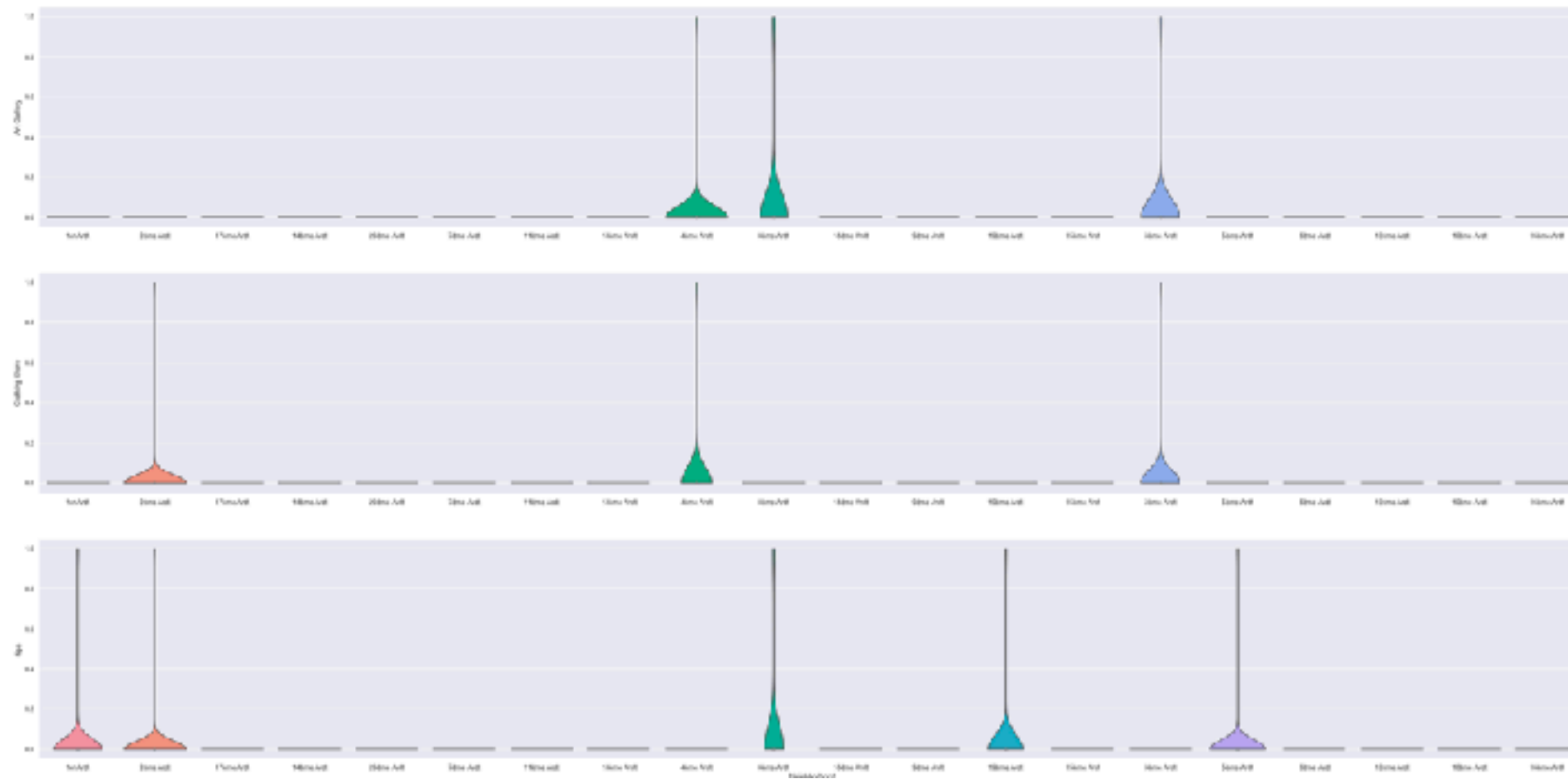
Battle of neighborhoods

Extract Data about Paris using Foursquare API

We are going to start utilizing the Foursquare API to explore the neighborhoods and segment them. We set the LIMIT parameter to **100**, which would limit the number of venues returned by the Foursquare API and the radius of 500 meter. Here is a head of the list of Paris City.

	name	categories	lat	lng
0	Musée du Louvre	Art Museum	48.860847	2.336440
1	Palais Royal	Historic Site	48.863236	2.337127
2	Comédie-Française	Theater	48.863088	2.336612
3	La Clef Louvre Paris	Hotel	48.863977	2.336140
4	Cour Napoléon	Plaza	48.861172	2.335088
5	Place du Palais Royal	Plaza	48.862523	2.336688
6	Vestige de la Forteresse du Louvre	Historic Site	48.861577	2.333608
7	Place Colette	Plaza	48.863145	2.335943
8	LouLou	Italian Restaurant	48.862804	2.333500
9	Les Arts Décoratifs	Art Museum	48.863077	2.333893
10	Mariage Frères	Tea Room	48.861437	2.334143
11	Jardin du Palais Royal	Garden	48.864941	2.337728
12	Cour Carrée du Louvre	Pedestrian Plaza	48.860360	2.335543
13	Kosyuen 萃儔園	Tea Room	48.864163	2.333567
14	La Vénus de Milo (Vénus de Milo)	Exhibit	48.859943	2.337234
15	Aesop	Cosmetics Shop	48.864158	2.333601
16	Cibus	Italian Restaurant	48.864757	2.335377
17	Astier de Villatte	Furniture / Home Store	48.864096	2.333808
18	Thémaé	Spa	48.863589	2.339756
19	Sanukiya	Udon Restaurant	48.864713	2.333805

Frequency distribution for the top 3 venue categories for each neighborhood (click to enlarge)



The business types criteria specified by the client (clothing stores, art galleries and spa)

These are the venue types that the client wants to have an abundant density of in the ideal Japanese restaurant locations. I've used a violin plot from the seaborn library - it is a great way to visualise frequency distribution datasets, they display a density estimation of the underlying distribution.

The Neighborhoods

So as we can see from the analysis there are 4 neighborhoods to open new restaurant - according to the criteria that they have the 2/3 specified venues in a great frequency (Clothing stores, spa, art gallery). They are as follows:

- 2eme Ardt
- 4eme Ardt
- 8eme Ardt
- 3eme Ardt

Battle of neighborhoods

The table on the right will show to us the neighborhoods where we have low concurence inorder to open our japanese restaurant.

The neighborhoods where we have low concurrence are :

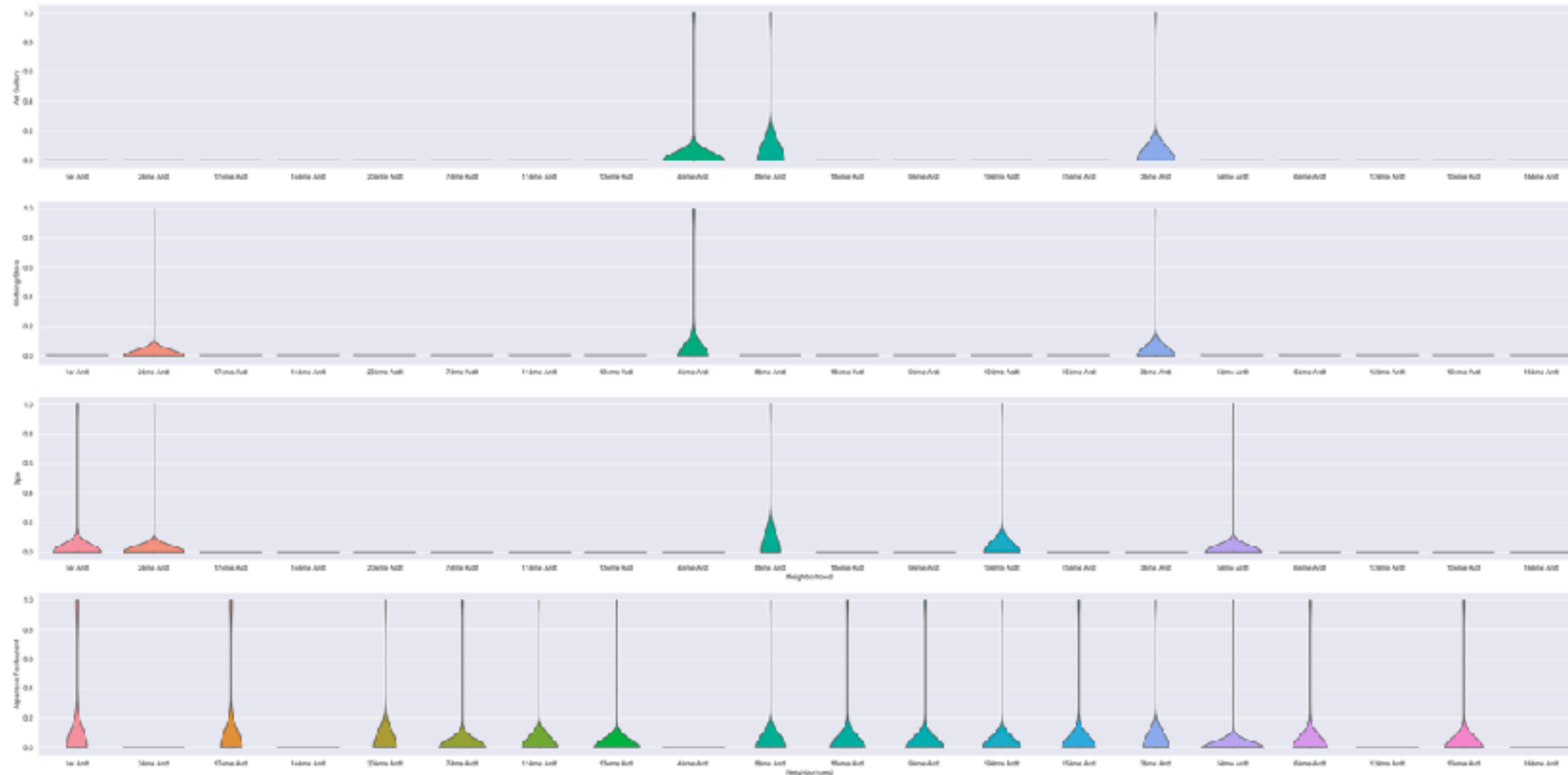
- 12eme Ardt
- 14eme Ardt
- 16eme Ardt
- 4eme Ardt
- 2eme Ardt

So if we take in to acount the neighborhoods where our criteria applies (Clothing store, spa, art gallery) we can say that the neighborhoods where we have low concurrence and where our criteria is respected are :

- 4eme Ardt
- 2eme Ardt

Japanese Restaurant		Neighborhood
2	0.0	12ème Ardt
4	0.0	14ème Ardt
6	0.0	16ème Ardt
14	0.0	4ème Ardt
12	0.0	2ème Ardt

Frequency distribution for the top 3 venue categories for each neighborhood (includes Japanese Restaurant)



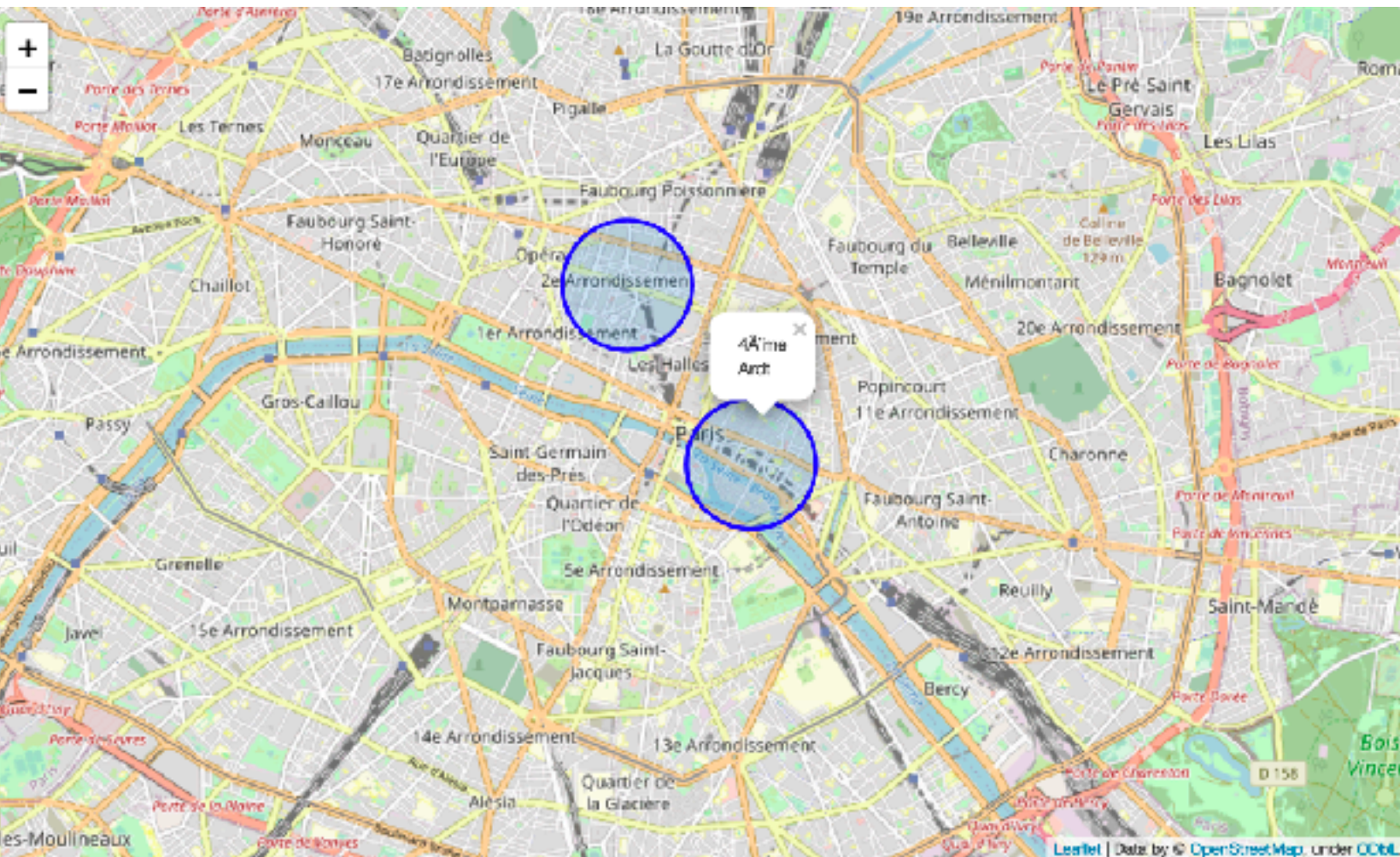
discussions

The reasoning being that if the 2/3 criteria are met - identifying lively neighbourhoods with art galleries, spas and clothing stores - the addition of japanese restaurant concurrence to the neighbourhood's mix of shops is an important bonus.

Thus, the last two potential new japanese restaurant locations are those where four criteria are met :

- 4eme Ardt
- 2eme Ardt

Conclusion



We have narrowed the search area down significantly from 20 potential districts to 2 that should suit the client's wish to open its restaurant.

Without the use of data to make targeted decisions, the process could have been lengthy and could have resulted in the new restaurant opening in non-standard areas for the customer. The data helped to provide a better strategy and a better way forward, and these data-driven decisions will lead to a better solution in the end.