# The Battle of Neighborhoods

# Report

# **Open restaurant in Paris**

# Contents

# 1. INTRODUCTION

## 1.1.Background

Paris is the capital and largest city of France. It has 2 million inhabitants and is the fifth of the most populated cities in the world. the city is one of the most influential centers worldwide, as well as a very strong cultural, political and economic one.

Paris is a city which enjoys an international influence not only from a tourist point of view but also from a professional point of view. This city is home to many business districts, each with their own particularities. Whether ultra-modern or more historic, the Parisian business districts form a veritable hive that comes alive every day. That's why, to start a bussniss in Paris shuch as opening a restaurent, we should search and analyse the location, to have our place in 12.000 restaurants and to be famous competitors.

## 1.2. Problem

Paris still always attractive for new business such as: opening new restaurant.

For this, we must make a detailed study to find the most suitable places and which still support a new restaurant. It depends on the geographic distribution of the competitors.

## 1.3.Interest

This project will be interesting for business man who wants to take a risk and start new adventure in fabulous city like Paris.

# 2. DATA

## 2.1. Data acquisition

The data used for this project comes from three different locations:

• Foursquare. It is a local search-and-discovery service which provides information on different types of entertainment, drinking and dining venues… Foursquare has an API that

can be used to query their database and find information related to the venues, such as location, overall category, reviews and tips.

• Paris Neighborhood Names and geographic coordinates. Available on https://opendata.paris.fr/explore/dataset/arrondissements/table/ , this is used to obtain the neighborhood location information from the city.

• Paris borough Names and geographic coordinates. Data available on https://opendata.paris.fr/explore/dataset/arrondissements/table/?dataChart

• List of Paris restaurant, were we can get the popular places, available in https://opendata.paris.fr/explore/dataset/restaurants-casvp/table/?disjunctive.code&disjunctive.nom_restaurant&disjunctive.type

In order to facilitate our work, we have based in one dataset downloadable from the below link:

https://raw.githubusercontent.com/AR-data-science/Coursera_Capstone/master/Arrondissements_.csv,

It presents the list of city's borough and their geographical coordination.

### 2.2. Data cleaning

During this step, we were invited to clean the dataset from unnecessary information, to get a dataset ready to be exploitable.

We will present the related work in the next section.

## 3. METHODOLOGY

### 3.1. Data preprocessing

- Neighborhood information's

During the data preprocessing stage, we prepare the data to be used during the clustering process.

| | CAR | NAME | NSQAR | CAR.1 | CARINSEE | LAR | NSQCO | SURFACE | PERIMETRE | Geometry_X | Geometry_Y |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 3 | Temple | 750000003 | 3 | 3 | 3eme Ardt | 750001537 | 1170882828 | 4519264 | 48.862872 | 2.360001 |
| 1 | 19 | Buttes-Chaumont | 750000019 | 19 | 19 | 19eme Ardt | 750001537 | 6792651129 | 11253182 | 48.887076 | 2.384821 |
| 2 | 14 | Observatoire | 750000014 | 14 | 14 | 14eme Ardt | 750001537 | 5614877309 | 10317483 | 48.829245 | 2.326542 |
| 3 | 10 | Entrepot | 750000010 | 10 | 10 | 10eme Ardt | 750001537 | 2891739442 | 6739375 | 48.876130 | 2.360728 |
| 4 | 12 | Reuilly | 750000012 | 12 | 12 | 12eme Ardt | 750001537 | 16314782637 | 24089666 | 48.834974 | 2.421325 |
| 5 | 16 | Passy | 750000016 | 16 | 16 | 16eme Ardt | 750001537 | 16372542129 | 17416110 | 48.860392 | 2.261971 |
| 6 | 11 | Popincourt | 750000011 | 11 | 11 | 11eme Ardt | 750001537 | 3665441552 | 8282012 | 48.859059 | 2.380058 |
| 7 | 2 | Bourse | 750000002 | 2 | 2 | 2eme Ardt | 750001537 | 991153745 | 4554104 | 48.868279 | 2.342803 |
| 8 | 4 | Hotel-de-Ville | 750000004 | 4 | 4 | 4eme Ardt | 750001537 | 1600585632 | 5420908 | 48.854341 | 2.357630 |
| 9 | 17 | Batignolles-Monceau | 750000017 | 17 | 17 | 17eme Ardt | 750001537 | 5668834504 | 10775580 | 48.887327 | 2.306777 |
| 10 | 18 | Buttes-Montmartre | 750000018 | 18 | 18 | 18eme Ardt | 750001537 | 5996051308 | 9916464 | 48.892569 | 2.348161 |
| 11 | 1 | Louvre | 750000001 | 1 | 1 | 1er Ardt | 750001537 | 1824612860 | 6054937 | 48.862563 | 2.336443 |
| 12 | 5 | Pantheon | 750000005 | 5 | 5 | 5eme Ardt | 750001537 | 2539374623 | 6239195 | 48.844443 | 2.350715 |
| 13 | 7 | Palais-Bourbon | 750000007 | 7 | 7 | 7eme Ardt | 750001537 | 4090057185 | 8099425 | 48.856174 | 2.312188 |
| 14 | 20 | Menilmontant | 750000020 | 20 | 20 | 20eme Ardt | 750001537 | 5983446037 | 10704940 | 48.863461 | 2.401188 |
| 15 | 8 | elysee | 750000008 | 8 | 8 | 8eme Ardt | 750001537 | 3880036397 | 7880533 | 48.872721 | 2.312554 |
| 16 | 9 | Opera | 750000009 | 9 | 9 | 9eme Ardt | 750001537 | 2178303275 | 6471588 | 48.877164 | 2.337458 |

**Figure 1. Initial borough's list**

To explore the above dataset, we should define the necessary information to keep.

| | borough_Num | borough | French_Name | Latitude | Longitude |
|---|---|---|---|---|---|
| 0 | 3 | Temple | 3eme Ardt | 48.862872 | 2.360001 |
| 1 | 19 | Buttes-Chaumont | 19eme Ardt | 48.887076 | 2.384821 |
| 2 | 14 | Observatoire | 14eme Ardt | 48.829245 | 2.326542 |
| 3 | 10 | Entrepot | 10eme Ardt | 48.876130 | 2.360728 |
| 4 | 12 | Reuilly | 12eme Ardt | 48.834974 | 2.421325 |
| 5 | 16 | Passy | 16eme Ardt | 48.860392 | 2.261971 |
| 6 | 11 | Popincourt | 11eme Ardt | 48.859059 | 2.380058 |
| 7 | 2 | Bourse | 2eme Ardt | 48.868279 | 2.342803 |
| 8 | 4 | Hotel-de-Ville | 4eme Ardt | 48.854341 | 2.357630 |
| 9 | 17 | Batignolles-Monceau | 17eme Ardt | 48.887327 | 2.306777 |
| 10 | 18 | Buttes-Montmartre | 18eme Ardt | 48.892569 | 2.348161 |
| 11 | 1 | Louvre | 1er Ardt | 48.862563 | 2.336443 |
| 12 | 5 | Pantheon | 5eme Ardt | 48.844443 | 2.350715 |
| 13 | 7 | Palais-Bourbon | 7eme Ardt | 48.856174 | 2.312188 |
| 14 | 20 | Menilmontant | 20eme Ardt | 48.863461 | 2.401188 |
| 15 | 8 | elysee | 8eme Ardt | 48.872721 | 2.312554 |
| 16 | 9 | Opera | 9eme Ardt | 48.877164 | 2.337458 |

**Figure 2. Dataset after renaming and removing columns**

At this stage, the dataset is ready to be exploitable.

### 3.2. Exploratory Data Analysis

- Foursquare API data

For this step, we used the Foursquare Places API. One of the features of this API is to provide a list of venues within a specific location, based on the Lat/Lon coordinates and a radius. In order to obtain a list of venues within a specified area, we use the "explore" endpoint from the API. By passing the proper parameters via an HTTP request to the explore endpoint, we get a JSON object with the information shown as below:

```
LIMIT = 100
radius = 500

url = 'https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&limit={}'.format(
    CLIENT_ID,
    CLIENT_SECRET,
    VERSION,
    borough_latitude,
    borough_longitude,
    radius,
    LIMIT)
```

```
# get the result
json_results = requests.get(url).json()
json_results
```

```
: {'meta': {'code': 200, 'requestId': '5e309b68c94979001b78c830'},
   'response': {'suggestedFilters': {'header': 'Tap to show:',
     'filters': [{'name': 'Open now', 'key': 'openNow'}]},
    'headerLocation': 'Palais-Royal',
    'headerFullLocation': 'Palais-Royal, Paris',
    'headerLocationGranularity': 'neighborhood',
    'totalResults': 142,
    'suggestedBounds': {'ne': {'lat': 48.86706270450001,
      'lng': 2.3432708805600218},
     'sw': {'lat': 48.8580626955, 'lng': 2.329615843439978}},
```

**Figure 3. Foursquare result**

By using a custom function, calling the "explore" endpoint, we created a dataset with the top 100 venues within 500 meters of the center of each neighborhood.

The location object contains the coordinates of each venue, which will be used to associate it with its respective neighborhood. The categories array will be used to categorize the neighborhood.

As below, we will present the list of the most famous venues in Paris and their longitude and latitude values.

| | 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 | Cour Napoléon | Plaza | 48.861172 | 2.335088 |
| 4 | Place du Palais Royal | Plaza | 48.862523 | 2.336688 |
| 5 | LouLou | Italian Restaurant | 48.862804 | 2.333500 |
| 6 | Place Colette | Plaza | 48.863145 | 2.335943 |
| 7 | Kodawari Ramen (Tsukiji) | Ramen Restaurant | 48.864383 | 2.336297 |
| 8 | Vestige de la Forteresse du Louvre | Historic Site | 48.861577 | 2.333508 |
| 9 | Grand Hôtel du Palais Royal | Hotel | 48.863183 | 2.337901 |
| 10 | Les Arts Décoratifs | Art Museum | 48.863077 | 2.333393 |
| 11 | Café Kitsuné | Coffee Shop | 48.863983 | 2.335794 |

**get total number of venue**

```
print('{} venues were returned by Foursquare.'.format(nearby_venues.shape[0]))
```

```
100 venues were returned by Foursquare.
```

**Figure 4. The 100 most famous places**

Basically, we will count how many venues from all available categories are found on each neighborhood, and then use that information to compare neighborhoods.

```
----10eme Ardt----
                venue  freq
0    French Restaurant  0.11
1              Bistro  0.05
2               Hotel  0.05
3         Coffee Shop  0.05
4                Café  0.04
5    Indian Restaurant  0.04
6         Pizza Place  0.03
7  Japanese Restaurant  0.03
8                 Bar  0.03
9   Italian Restaurant  0.03
----11eme Ardt----
                venue  freq
0    French Restaurant  0.10
1                Café  0.07
2            Wine Bar  0.04
3          Restaurant  0.04
4         Pastry Shop  0.04
```

**Figure 5. Venues frequency / borough**

To be more precise, we will get top 10 venues for each borough.

| | borough | 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 | Supermarket | Sandwich Place | Zoo Exhibit | Electronics Store | Food & Drink Shop | Flower Shop | Fish & Chips Shop | Fast Food Restaurant | Farmers Market | Falafel Restaurant |
| 1 | 10eme Ardt | French Restaurant | Bistro | Hotel | Coffee Shop | Café | Indian Restaurant | Pizza Place | Japanese Restaurant | Italian Restaurant | Bar |
| 2 | 11eme Ardt | French Restaurant | Café | Pastry Shop | Wine Bar | Restaurant | Italian Restaurant | Cocktail Bar | Pizza Place | Theater | Grocery Store |
| 3 | 12eme Ardt | Zoo Exhibit | Zoo | Monument / Landmark | Supermarket | Bike Rental / Bike Share | Park | Ethiopian Restaurant | Food & Drink Shop | Flower Shop | Fish & Chips Shop |
| 4 | 13eme Ardt | Vietnamese Restaurant | Asian Restaurant | Thai Restaurant | Chinese Restaurant | French Restaurant | Juice Bar | Japanese Restaurant | Furniture / Home Store | Cambodian Restaurant | Fast Food Restaurant |
| 5 | 14eme Ardt | French Restaurant | Hotel | Bistro | Supermarket | Bakery | Brasserie | Fast Food Restaurant | Sushi Restaurant | Tea Room | Food & Drink Shop |
| 6 | 15eme Ardt | Hotel | Italian Restaurant | French Restaurant | Coffee Shop | Thai Restaurant | Park | Japanese Restaurant | Indian Restaurant | Lebanese Restaurant | Bakery |
| 7 | 16eme Ardt | Lake | Plaza | Recording Studio | Park | Bus Stop | Bus Station | French Restaurant | Boat or Ferry | Pool | Art Museum |
| 8 | 17eme Ardt | French Restaurant | Hotel | Italian Restaurant | Japanese Restaurant | Plaza | Bakery | Café | Bistro | Restaurant | Burger Joint |

**Figure 6. Top 10 venues / neighborhoods**

## 3.3. Clustering

We have transformed the above presentation into a violin plot from the seaborn library - it is a great way to visualize frequency distribution datasets, they display a density estimation of the underlying distribution.

I have picked 11 venues categories related to various restaurant (French, Lebanese, Italian, Japanese, Chinese….)

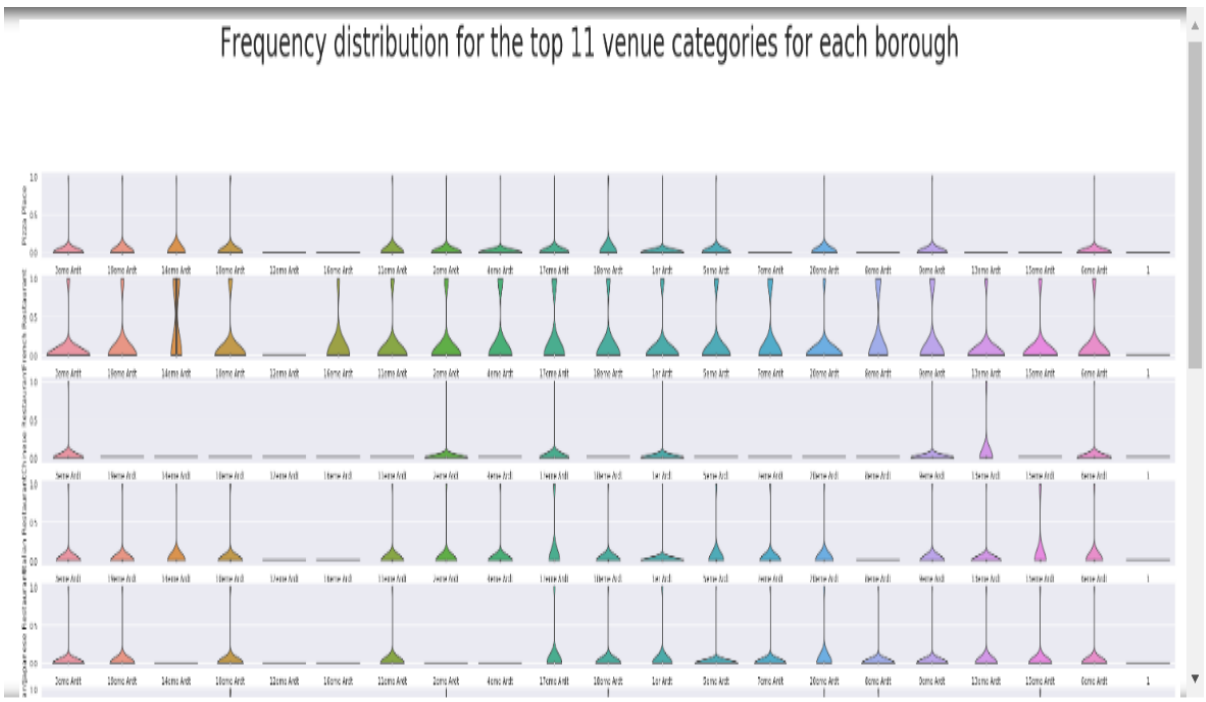


**Figure 7. Frequency distibution for 11 venue categories**

Now, the task is easier and clearer: we will pick the neighborhood having low density of restaurant.

## 4. RESULTS

After clustering the Paris borough based on the results from the Foursquare API data, we were able to select the best 3 neighborhoods to open in our business.

| | Arrondissement_Num | Borough | French_name | Latitude | Longitude |
|---|---|---|---|---|---|
| **0** | 12 | Reuilly | 12ème | 48.835 | 2.42132 |
| **1** | 16 | Passy | 16ème | 48.8604 | 2.26197 |
| **2** | 20 | Menilmontant | 20ème | 48.8635 | 2.40119 |

**Figure 8. Result list**

Map representation for selected neighborhoods:


**Figure 9. Map Representation**

## 5. DISCUSSION

The selected borough has similar characteristics. Most of them are dominated by museum, historical site, garden, shopping mall… are densely populated borough and have few types of restaurants.

This constitutes good options for opening a restaurant.

## 6. CONCLUSSION

We were able to determine a good set to open a new restaurant, considering the variables described in the previous sections.

During this project I applied several methodologies used during the course, such as data wrangling with pandas, basic data visualization and Foursquare API.

For future projects with similar characteristics, it should be considered to expand the amount of data available (for example, using the premium features of the Foursquare API).