# Capstone Project

The Battle of  Neighborhoods

Paris Restaurants

Lucas DA RIN

# Agenda

## Introduction

**The Battle of Neighborhoods**

- Choice of theme
- Neighborhoods in Paris
- What's the goal of this Project

## Data

**Data to be used**

- Sources of data
- Foursquare API
- Challenges of this Project

**Data Cleaning**

## Methodology

**Obtaining Data**

- The table of neighborhoods
- Functions
- Challenges of this Project

**Data Cleaning**

## Results & Discussion

**Results**

- Best and Most Popular Restaurants
- Study of the neighborhoods

**Discussion**

- Do the answers make sense?
- Limitations of the model

## Conclusion

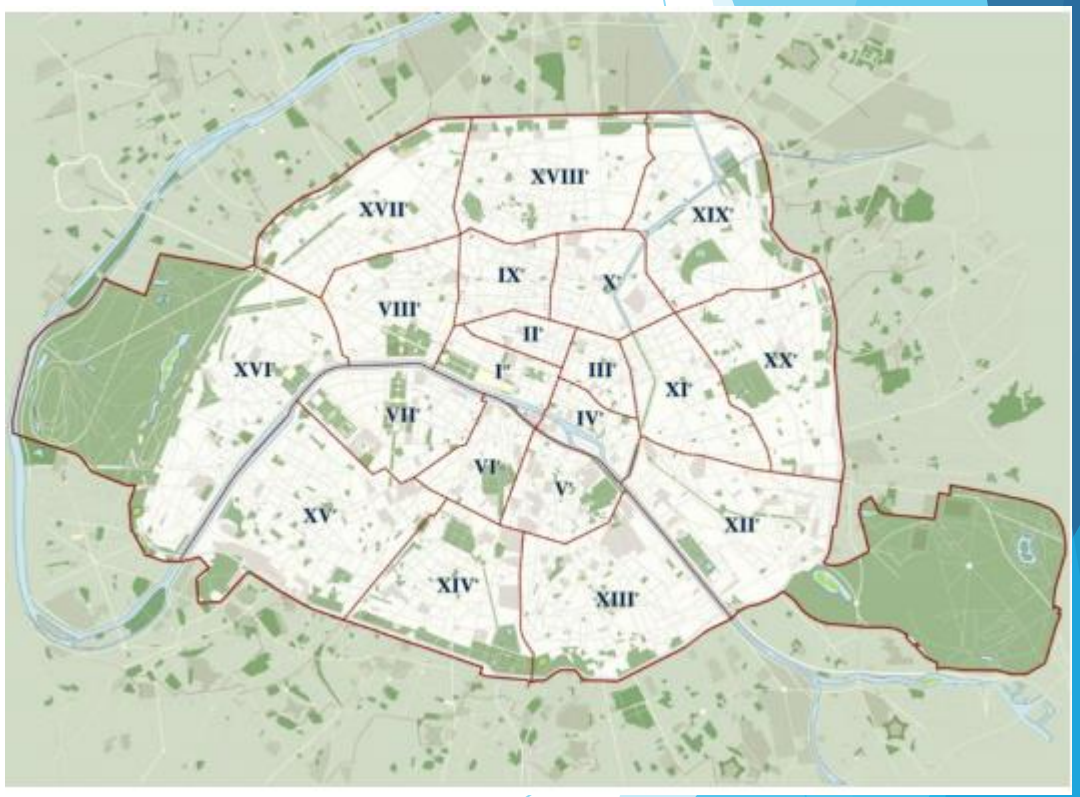
**Questions answered**

- Are french restaurants the most common ones in Paris?
- What are best neighborhoods from the culinary point of view?
- Which neighborhoods

# Introduction

- Paris is one of the most important cities of the world when it comes to tourism, arts, fashion and culinary. This Project focuses on the culinary of the city;

- Paris has 20 neighborhoods and the goal of this Project is examining how restaurants are distributed among them



# Data

## Data Sources

- The data about the neighborhoods of Paris comes from the French Government (https://www.data.gouv.fr/);
- The data contains geographic locations for each of the 20 neighborhoods of the city;
- Some of the columns are not used in this Project.

## Foursquare API

- Foursquare API is the tool for finding venues given locations;
- In this Project, the API is used for finding venues near the locations provided in the table;
- Later on this Project the API is used to find clients' opinions about the restaurants;

# Data Cleaning

| | 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 |
| 17 | 13 | Gobelins | 750000013 | 13 | 13 | 13eme Ardt | 750001537 | 7149311091 | 11546547 | 48.828388 | 2.362272 |
| 18 | 15 | Vaugirard | 750000015 | 15 | 15 | 15eme Ardt | 750001537 | 8494994081 | 13678798 | 48.840085 | 2.292826 |
| 19 | 6 | Luxembourg | 750000006 | 6 | 6 | 6eme Ardt | 750001537 | 2153095586 | 6483687 | 48.849130 | 2.332898 |

Many of the columns in the table on the left, provided by our source, are not used in this Project. For this reason some of them were dropped to arrive at the table form below, with the columns needed

| | ZIP Code | Neighborhood | SURFACE | Latitude | Longitude |
|---|---|---|---|---|---|
| 11 | 75001 | 1er Ardt | 1824612860 | 48.862563 | 2.336443 |
| 7 | 75002 | 2eme Ardt | 991153745 | 48.868279 | 2.342803 |
| 0 | 75003 | 3eme Ardt | 1170882828 | 48.862872 | 2.360001 |
| 8 | 75004 | 4eme Ardt | 1600585632 | 48.854341 | 2.357630 |
| 12 | 75005 | 5eme Ardt | 2539374623 | 48.844443 | 2.350715 |

# Methodology

- This project consists of exploring venues around specific neighborhoods, using Foursquare API, investigating the ratings and popularity of venues and exploring the distribution of venues and their ratings across neighborhoods;
- The venues were found after the definition of a function for finding 100 results in a radius of, at first, 500 meters, as shown on the right.

```
def getNearbyVenues(zipcodes, names, latitudes, longitudes):
    radius=500
    LIMIT=100
    venues_list=[]
    for zipc, name, lat, lng in zip(zipcodes, names, latitudes, longitudes):
        print(name)

        # create the API request URL
        url = 'https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&limit={}'.fo
            CLIENT_ID,
            CLIENT_SECRET,
            VERSION,
            lat,
            lng,
            radius,
            LIMIT)

        # make the GET request
        results = requests.get(url).json()["response"]['groups'][0]['items']

        # return only relevant information for each nearby venue
        venues_list.append([(
            zipc,
            name,
            lat,
            lng,
            v['venue']['id'],
            v['venue']['name'],
            v['venue']['location']['lat'],
            v['venue']['location']['lng'],
            v['venue']['categories'][0]['name']) for v in results])

    nearby_venues = pd.DataFrame([item for venue_list in venues_list for item in venue_list])
    nearby_venues.columns = ['ZIP Code',
                'Neighborhood',
                  'Neighborhood Latitude',
                  'Neighborhood Longitude',
                'Id',
                  'Venue',
                  'Venue Latitude',
                  'Venue Longitude',
                  'Venue Category']

    return(nearby_venues)
```

# Methodology

- The data about the neighborhoods comes in a csv(comma-separated values) file, consisting of the ZIP Code, Neighborhood Name and Geographical locations;
- The venues found were treated in order to keep the most likely of the occurrences in cases of duplicates;
- The data of the venues was then treated in order to keep only restaurants and bars.

| | Category | count |
|---|---|---|
| 0 | French Restaurant | 354 |
| 1 | Hotel | 169 |
| 2 | Italian Restaurant | 104 |
| 3 | Bar | 82 |
| 4 | Bakery | 76 |
| 5 | Japanese Restaurant | 69 |
| 6 | Bistro | 65 |
| 7 | Plaza | 62 |
| 8 | Café | 54 |
| 9 | Coffee Shop | 52 |
| 10 | Wine Bar | 46 |
| 11 | Restaurant | 40 |
| 12 | Cocktail Bar | 38 |
| 13 | Pizza Place | 36 |
| 14 | Vietnamese Restaurant | 35 |
| 15 | Thai Restaurant | 30 |
| 16 | Garden | 28 |
| 17 | Supermarket | 28 |
| 18 | Park | 27 |
| 19 | Asian Restaurant | 26 |
| 20 | Pastry Shop | 22 |
| 21 | Seafood Restaurant | 20 |
| 22 | Art Museum | 20 |
| 23 | Bookstore | 20 |
| 24 | Art Gallery | 20 |
| 25 | Ice Cream Shop | 19 |
| 26 | Dessert Shop | 19 |
| 27 | Creperie | 19 |
| 28 | Sandwich Place | 18 |
| 29 | Cheese Shop | 17 |
| 30 | Clothing Store | 17 |

List of types of venues

# Methodology – Filtering Restaurants & Bars

- The criteria chosen was to check for 35 most common kinds of venues and filtering out those that were not related to restaurants and bars or that could not fit in “French Restaurants”, “Italian Restaurants”, “Cafés & Bars” and “Asian Restaurants” categories;

- The list consists of the following types of venues ['French Restaurant','Italian Restaurant', 'Japanese Restaurant', 'Bistro', 'Wine Bar','Coffee Shop','Café','Cocktail Bar','Bar','Vietnamese Restaurant','Asian Restaurant', 'Thai Restaurant', 'Creperie', 'Chinese Restaurant' ];

- 989 venues were found.

# Methodology - ‘Rating’, ‘Likes’ and ‘Tips’

- A function using Foursquare API was defined for extracting ‘Rating’, ‘Likes’ and ‘Tips’ fields;
- The merging of the venues details table and the entire list of venues was then used for the analysis of distribution or rates and geographical distribution of the venues;
- The results and discussion of the investigation are presented in the ‘Results’ and ‘Discussion’ sections.

```
def get_venue_details(venue_id):
    venue_details = []

    for ven in venue_id:
    #url to fetch data from foursquare api

        url = 'https://api.foursquare.com/v2/venues/{}?&client_id={}&client_secret={}&v={}'.format(
                ven,
                CLIENT_ID,
                CLIENT_SECRET,
                VERSION)
        # get all the data
        results = requests.get(url).json()
        print(results)
        venue_data = results["response"]

        try:
            venue_id=venue_data['venue']['id']
            venue_name=venue_data['venue']['name']
            venue_likes=venue_data['venue']['likes']['count']
            venue_rating=venue_data['venue']['rating']
            venue_tips=venue_data['venue']['tips']['count']
            venue_details.append([venue_id,venue_name,venue_likes,venue_rating,venue_tips])
        except KeyError:
            pass
    column_names=['ID','Name','Likes','Rating','Tips']
    df = pd.DataFrame(venue_details,columns=column_names)
    return df
```

# Results

- The results show that French Restaurants are the most popular ones in 11 of the 20 neighborhoods of Paris, while Cafés & bars and Asian Restaurants account for 7 and 2 neighborhoods, respectively;

- Café de Flore is the most popular one and La Cave de Septime is the best restaurant of our sample;

- 3ème neighborhood is the one with the best ratings, while 12ème and 20ème have average ratings below 7;

| | ZIP Code | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue |
|---|---|---|---|---|
| 0 | 75010 | Cafés and Bars | French Restaurant | Asian Restaurants |
| 1 | 75011 | Cafés and Bars | French Restaurant | Italian Restaurants |
| 2 | 75012 | French Restaurant | Cafés and Bars | Italian Restaurants |
| 3 | 75013 | Asian Restaurants | French Restaurant | Cafés and Bars |
| 4 | 75014 | French Restaurant | Cafés and Bars | Asian Restaurants |
| 5 | 75015 | French Restaurant | Cafés and Bars | Italian Restaurants |
| 6 | 75016 | French Restaurant | Italian Restaurants | Asian Restaurants |
| 7 | 75017 | French Restaurant | Italian Restaurants | Cafés and Bars |
| 8 | 75018 | Cafés and Bars | French Restaurant | Asian Restaurants |
| 9 | 75019 | Cafés and Bars | French Restaurant | Asian Restaurants |
| 10 | 75001 | Asian Restaurants | French Restaurant | Italian Restaurants |
| 11 | 75020 | Cafés and Bars | French Restaurant | Asian Restaurants |
| 12 | 75002 | Cafés and Bars | French Restaurant | Italian Restaurants |
| 13 | 75003 | Cafés and Bars | French Restaurant | Asian Restaurants |
| 14 | 75004 | French Restaurant | Cafés and Bars | Italian Restaurants |
| 15 | 75005 | French Restaurant | Cafés and Bars | Asian Restaurants |
| 16 | 75006 | French Restaurant | Cafés and Bars | Italian Restaurants |
| 17 | 75007 | French Restaurant | Cafés and Bars | Italian Restaurants |
| 18 | 75008 | French Restaurant | Cafés and Bars | Asian Restaurants |
| 19 | 75009 | French Restaurant | Cafés and Bars | Italian Restaurants |

# Discussion

- The results show that French restaurants, followed by cafés and bars are the most common venues in 18 of the 20 neighborhoods;
- This first result is no surprise, since French restaurants are probably the most common ones in France;
- On the other hand, the 1st and the 13th neighborhoods have more Asian Restaurants than French and Bars, which makes sense, since the Japanese area and the Asian area are located in these neighborhoods respectively;

Quartier Asiatique. Source: Le Routard (routard.com)

# Discussion – Limitations of the model

- The model created for this project uses a sample of the venues found in Paris and it does not intend to represent with perfect accuracy the reality;
- The analysis of the best restaurant and the most popular one may be somehow compromised, since not all of the restaurants are in our sample;
- The number of results in the sample of the project is limited by the API quota, the number of restaurants and bars, 989, is the best one to be reached within the limits of the API quota;

# Conclusion

- This study concludes that french restaurants are most common ones in 11 of the 20 neighborhoods of Paris, followed by cafés and bars;
- Another interesting point is that the 13th and the 1st shows more asian restaurants than any category of restaurant in our sample;
- Another interesting conclusion is that most neighborhoods have no significant difference in their average rating, while 12ème and 20ème classify by far as the least appealing neighborhoods in clients' opinion.