

# OPENING A RESTAURANT IN PARIS

SIMON D.

10/03/2019

## I. Introduction

According to Insee [www.insee.fr](http://www.insee.fr), there are more than 45,000 restaurants in Paris and about 2.2 million people. London has approximately 16,000 restaurants for 9 million people. That is why opening a new restaurant in Paris can be an extremely challenging task compared to doing so in London. Choosing a restaurant type and a good spot, an entrepreneur or an investor usually relies on common sense and domain knowledge. Needless to say that too often an inconsiderate decision leads to a poor income and inevitable bankruptcy.

According to several surveys, up to 40% of such start-ups fail in the very first year. Let's suppose, an investor has enough time and money, as well as a passion to open the best eating spot in Paris. What type of restaurant would it be? What would be the best place for it?

What if there is a way to cluster city neighborhoods, based on their near-by restaurant similarity? What if we can visualize these clusters on a map? What if we might find what type of restaurant is the most and least popular in each location? Equipped with that knowledge, we might be able to make a smart choice from a huge number of restaurant types and available places. Let us allow machine learning to get the job done.

Target audience: investors, entrepreneurs, and chefs interested in opening a restaurant in Paris, who may need a piece of objective advice of what type of restaurant would be more successful and where exactly it should be opened.

## II. Data

This project will be using data from Paris open source database <https://opendata.paris.fr>, to collect information about boroughs and neighborhoods and merge those two separate dataframes.

We will also use Foursquare API [www.foursquare.com](http://www.foursquare.com) to collect the top 100 restaurants for each location

## III. Methodology

1. Using two tables from Paris open source database, collect information about Paris boroughs and neighborhoods.
2. Merge these two separate dataframes into one which will be used for the next steps.
3. Use the Geopy and Folium library to get the coordinates of every locations and map geospatial data on a Paris map.
4. Using Foursquare API, collect the top 100 restaurants and their categories for each location within a radius 300 meters.
5. Group collected restaurants by location and by taking the mean of the frequency of occurrence of each type, preparing them for clustering.
6. Cluster restaurants by k-means algorithm and analyze the top 10 most common restaurants in each cluster.
7. Visualize clusters on the map, thus showing the best locations for opening the chosen restaurant.

## IV. Exploring Paris Restaurants

We used Foursquare API to get top 100 restaurants in each neighborhood and we obtained 4791 restaurants of 123 individual types.

We applied one-hot encoding and grouped them by taking the mean of the frequency of occurrence of each type.

We clustered restaurants using the k-means algorithm based on their similarity.

## V. Results

### Cluster #1

The most common restaurant : French Restaurant

The 10th most common restaurant : Café

### Cluster #2

The most common restaurant : French Restaurant

The 10th most common restaurant : Diner

### Cluster #3

The most common restaurant : French Restaurant

The 10th most common restaurant : Sushi Restaurant

### Cluster #4

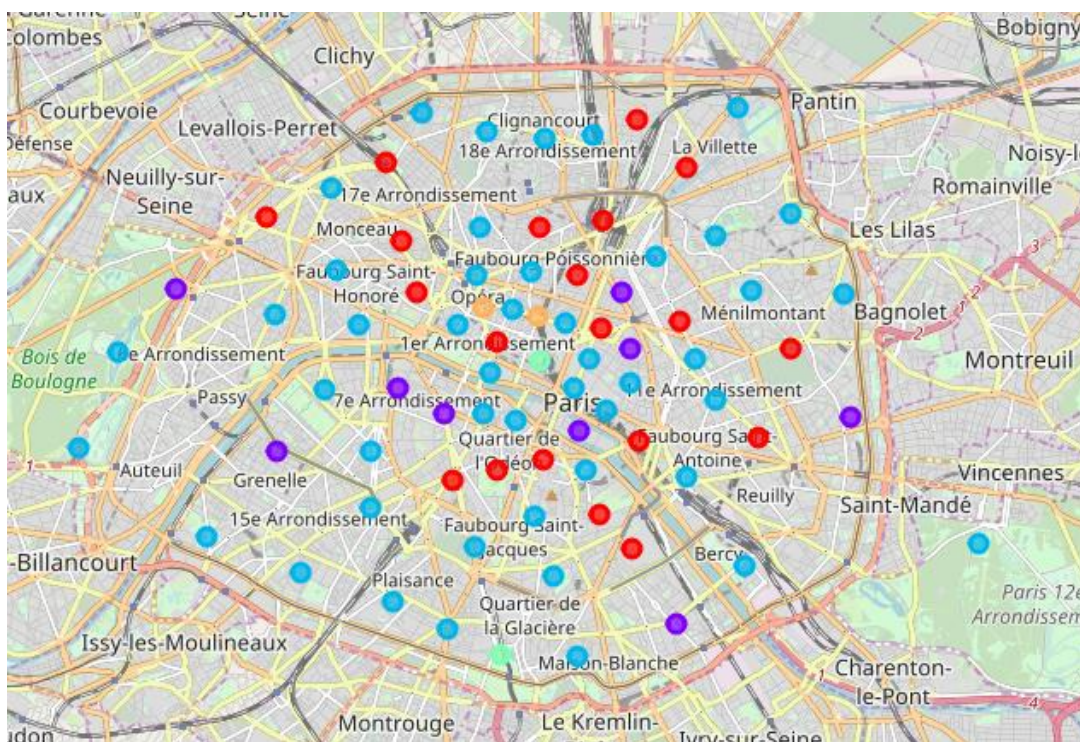
The most common restaurant : French Restaurant

The 10th most common restaurant : Diner

### Cluster #5

The most common restaurant : French Restaurant

The 10th most common restaurant : Burger Joint



#### MAP LEGEND

Cluster 1 - Red dots

Cluster 2 - Purple dots

Cluster 3 - Blue dots

Cluster 4 - Green dots

Cluster 5 - Orange dots

## VI. Discussion

Analyzing the most popular restaurants in each cluster, the stakeholder should select the least popular types within the top 10 as a safe choice.

In our recommendations, we advise selecting from the 9th and 10th positions as a reasonable balance between being too popular and having no customers.

Recommendations, based on description of each cluster:

**Cluster 1 Locations: Japanese Restaurant or Café**

**Cluster 2 Locations: Middle Eastern Restaurant or Diner**

**Cluster 3 Locations: Pizza Place or Sushi Restaurant**

**Cluster 4 Locations: French Restaurant or Diner**

**Cluster 5 Locations: Ramen Restaurant or Burger Joint**

## VII. Conclusion

In this report we established a methodology to determine what the most promising type of restaurant is , and where it should be opened.

This type of analysis can be applied to any city of your choice that has available geospatial information.

This type of analysis can be applied to any type of venue (shopping, clubs, etc.) that is available in Foursquare database.