

# APPLIED DATASCIENCE CAPSTONE FINAL ASSIGNMENT REPORT

## Business Problem

For this capstone project, I came up with an idea to leverage the Foursquare location data to explore the city of Paris, where I live. I love my city and i love eat japanease food especially sushy. So i'm going to work about this topic.

In this project we will try to find an optimal location for opening an Japanease restaurant in Paris. Since there are lots of japanease restaurants in Paris we will try to detect **locations with no Japanease restaurants** in vicinity. We would also prefer **locations as close to city center as possible**.

Let's meet these two conditions !

Based on definition of our problem, factors that will influence our decission are:

- number of existing restaurants in the neighborhood (any type of restaurant)
- number of and distance to Japanease – Sushi restaurants in the neighborhood, if any distance of neighborhood from city center We decided to use regularly spaced grid of locations, centered around city center, to define our neighborhoods.

Answering this problem would allow businessman and future restorers make sure that they do not miss the ideal place to set up a business in Paris !

## Data

Following data sources will be needed to extract/generate the required information:

centers of candidate areas will be generated algorithmically and approximate addresses of centers of those areas will be obtained using Geopy reverse geocoding number of restaurants and their type and location in every neighborhood will be obtained using Foursquare API coordinate of Paris center will be obtained using Geopy geocoding of well known Paris location (Chatelet les halles => Eglise Saint Eustache)

## Methodology

Let me remind you our problem : Find the optimal location for a sushi shop restaurant that's mean we search an empty place to set up our business.

To do this, let's create a grid of area candidates, equally spaced, centered around city center (lat and long are obtained using geopy) and within ~4,5km from Eglise Saint Eustache. Our neighborhoods will be defined as circular areas with a radius of 300 meters, so our neighborhood centers will be 600 meters apart.

To accurately calculate distances we need to create our grid of locations in Cartesian 2D coordinate system which allows us to calculate distances in meters (not in latitude/longitude degrees). Then we'll project those coordinates back to latitude/longitude degrees to be shown on Folium map.

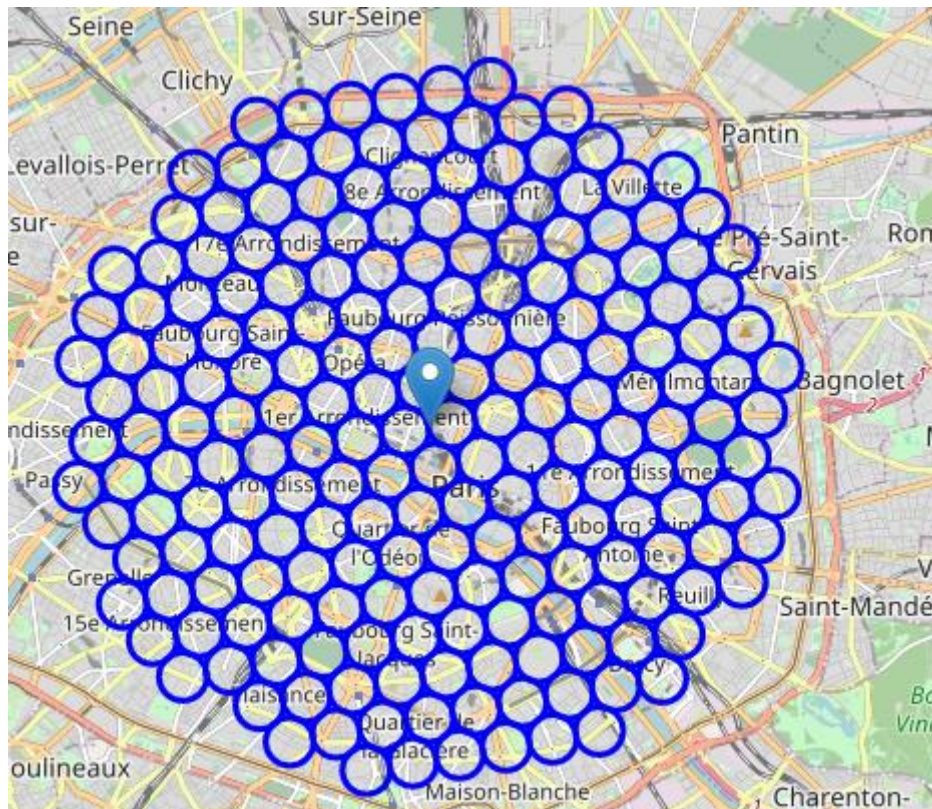


Fig 1 : Grid of potential neighborhoods

After obtaining our grid, we realise a dataframe with a new information : the estimated adress of those locations

	Address	Latitude	Longitude	X	Y	Distance from center
0	Réservoir de Montsouris, Avenue René Coty, Qua...	48.824688	2.334530	-428573.621169	5.485813e+06	4419.275959
1	49, Rue Boussingault, Cité Florale, Quartier d...	48.825579	2.342506	-427973.621169	5.485813e+06	4253.234064
2	7, Rue Martin Bernard, Cité Florale, Quartier ...	48.826469	2.350482	-427373.621169	5.485813e+06	4167.733197
3	4, Rue Toussaint-Féron, Cité Florale, Paris 13...	48.827358	2.358458	-426773.621169	5.485813e+06	4167.733197
4	Crèche Collective Municipale, Rue Sthrau, Pari...	48.828247	2.366435	-426173.621169	5.485813e+06	4253.234064
5	Restaurant universitaire CROUS Tolbiac, Rue Le...	48.829135	2.374412	-425573.621169	5.485813e+06	4419.275959
6	45, Rue des Plantes, Quartier du Petit-Montrou...	48.827912	2.321398	-429473.621169	5.486333e+06	4357.751714
7	26, Rue du Commandeur, Quartier du Petit-Montr...	48.828803	2.329374	-428873.621169	5.486333e+06	4058.324778
8	GHU Paris Psychiatrie & neurosciences - site S...	48.829694	2.337350	-428273.621169	5.486333e+06	3830.143600
9	Quartier de Croulebarbe, Paris, Île-de-France,...	48.830584	2.345326	-427673.621169	5.486333e+06	3686.461718

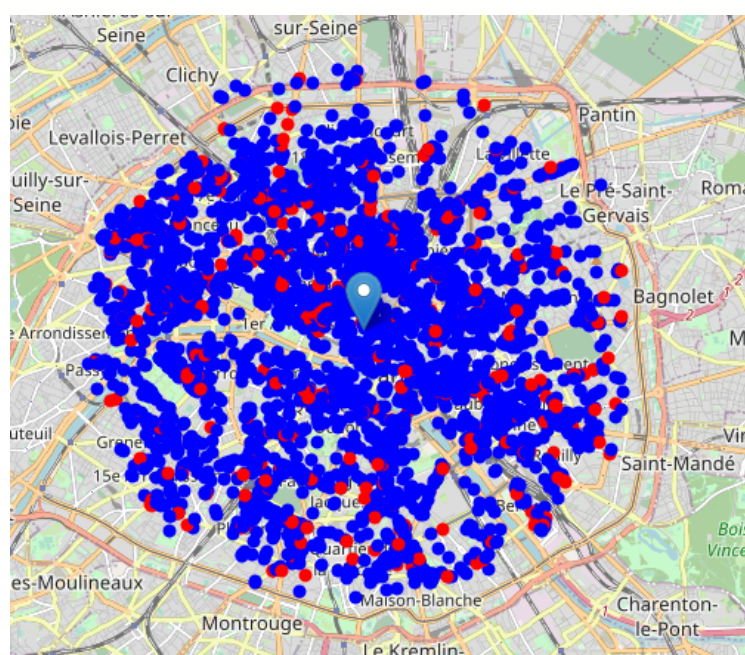
Fig2 : our dataframe pandas

Now that we have our location candidates, let's use Foursquare API to get info on restaurants in each neighborhood.

Coffe shops, pizza places, bakeries etc. are not direct competitors so we don't care about those. We will include in our list only venues that have 'restaurant' in category name, and we'll make sure to detect and include all the subcategories of specific 'japanese – sushi restaurant' category, as we need info on Japanese - sushi restaurants in the neighborhood. Roots were taken from Foursquare web site

General restaurant code = '4d4b7105d754a06374d81259' and sushi restaurant code = '4bf58dd8d48988d111941735',

We will obtain addresses of each restaurant in each area. Using the Folium library, we create a map of Paris showing the places of restaurants and japanese restaurants.

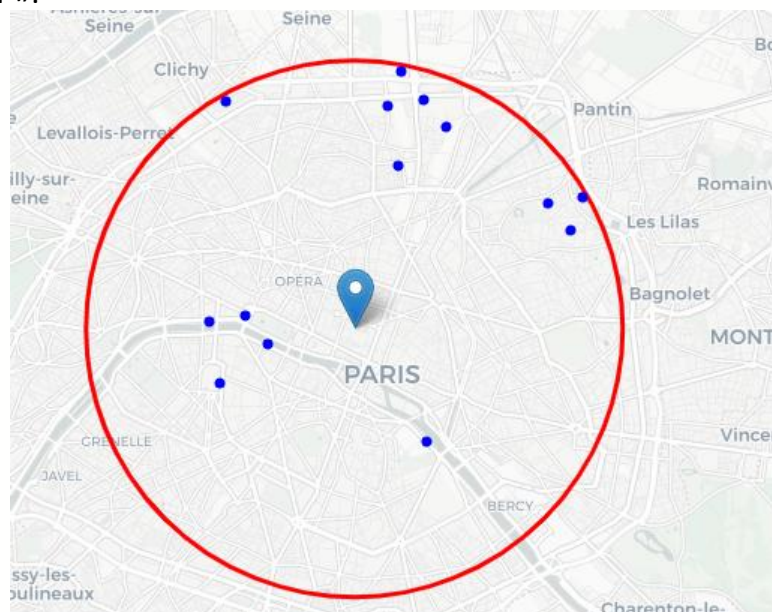




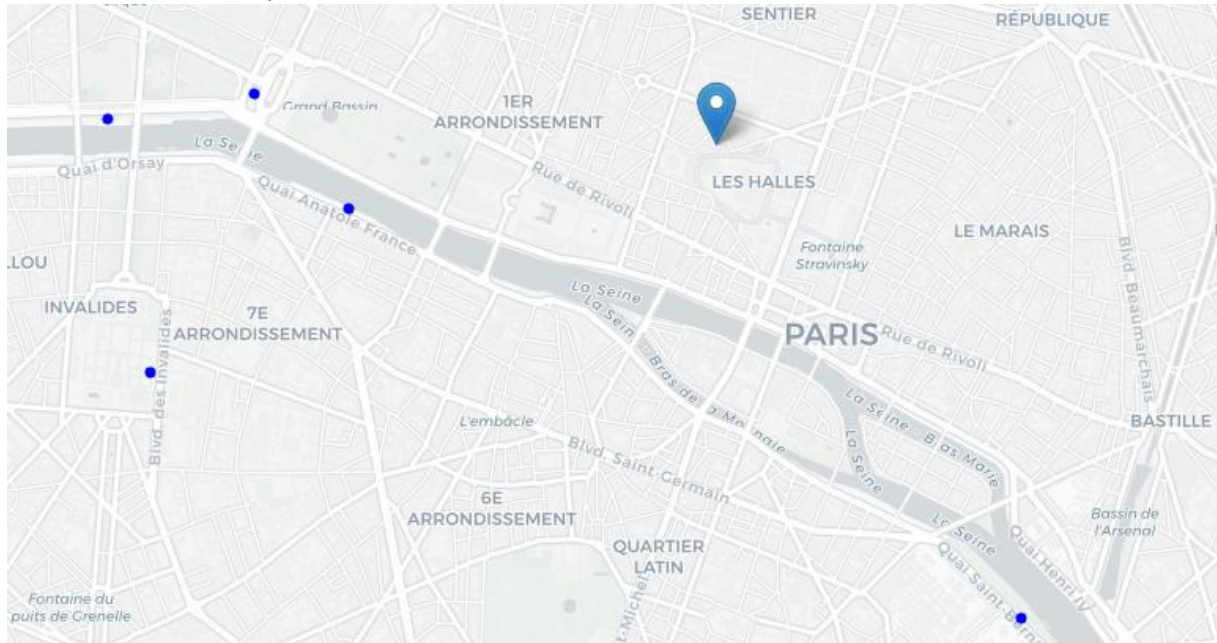
Second step in our analysis will be calculation and exploration of 'restaurant density' across different areas of paris - we will use heatmaps to identify a few promising areas close to center with low number of restaurants in general (and no Japanese restaurants in vicinity) and focus our attention on those areas. Unfortunately, the result don't looks like relevant because paris is very dense area and this view is not precise.



In final step we will take into consideration locations with no more than two restaurants in radius of 250 meters, and we want locations without Japanese restaurants in radius of 400 meters. We will get only few candidates and after a zoom analyze we determine one ideal place. Indeed, we eliminate areas unpleasant near « périphérique parisien » and « quartier de la goutte d'or ».



Our final candidates/ locations are 5.



## Results

Following our study, if business would like to set up a sushi shop in Paris, we would advise them to go mainly to the center of the capital, and precisely to settle down according our 5 areas. For example Quai saint bernard.

## Conclusion

To conclude, this study was an opportunity to highlight the ideal location(s) for a specific condition in a specific area. Thanks to the Foursquare API and the data available on the Geopy/ Google maps API, we were able to establish by neighborhood the number of restaurants in paris and their places, for each of them, in order to determine empty places.