OPENING A NEW JAPANESE RESTAURANT IN PARIS

APPLIED DATA SCIENCE CAPSTONE BY IBM - COURSERA MOOC

APRIL 2020

PARIS

Paris is the capital of France:

- 105 km^2
- 2.150 millions of people
- About 11700 restaurants

For many people, Paris is the most beatiful city of the World. At least, it's the most visited one. If museum, art gallery and historical buildings are the real reasons to visit, everyone has to lunch.

Moreover, Paris is the economical heart of France and about 10% of French people are working in Paris or around. And they have to lunch.

JAPANESE FOOD

France is well known for its cooking traditions. A lot of French restaurants are opened in Paris.

But Japanese food is becoming increasingly popular. So my project is to find the best place in Paris to install a new japanese restaurant.

DATA AND PROCESS

- Download the position of each of the 20 boroughs of Paris
- Download the 100 first venues in the 500m around the position of each borough
- Make some analysis on the datas
- Use *K-Means Clustering* to classify
- Find the right cluster to install the new restaurant :
  - not too much japanese restaurants
  - some restaurants to define a cluster where people are going to restaurants.

ANALYSIS

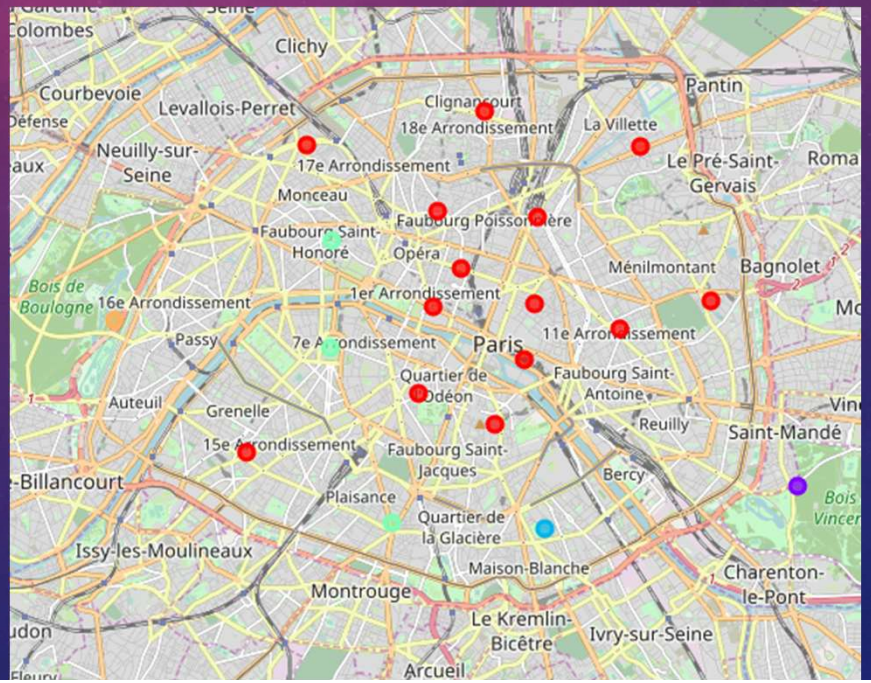
Using the top 100 downloaded for **Foursquare**, the first task is to calculate the nomber of category.

Then how many venues of each category is an relevant data, making possible to calculate the percentage of each category from all the venues of the borough.

These data are the basis to apply a clusterization

CLUSTERIZATION

- Using *k-means* algorithms, the venues will permit to clusterized the boroughs.
- 5 clusters seems to be the good **k** : 6 and 7 don't give more specific information, 4 is not enough.



LIMITATIONS AND SUGGESTIONS FOR FUTURE RESEARCH

In this project, only one factor is used : the frequency of occurrence on **Foursquare**. There are other factors such as population and income of residents that could influence the location decision of a new restaurant.

However, the kind of restaurants and shops around gives some clues, and there's no more precise data available at the time of this project.

Future research should look for a methodology to estimate the other factors, and use other information, using paid account to **Foursquare** API for example, to bypass limitations and be more precise to locate the right place.

RESULTS

The right cluster seems to be the 4th one, so one of this 3 boroughs :

- 7th arrondissement : **Palais-Bourbon**
- 14th arrondissement : **Observatoire**
- 8th arrondissement : **Elysée**

	Number	Name	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
6	7	Palais-Bourbon	Hotel	French Restaurant	Italian Restaurant	Café	Plaza	Bistro	Cocktail Bar	History Museum	Art Museum	Gourmet Shop
13	14	Observatoire	French Restaurant	Hotel	Supermarket	Pizza Place	Convenience Store	Italian Restaurant	Food & Drink Shop	Brasserie	Bakery	Bistro
7	8	Élysée	French Restaurant	Hotel	Spa	Cocktail Bar	Mediterranean Restaurant	Theater	Modern European Restaurant	Corsican Restaurant	Sporting Goods Shop	Resort

MERCI!

HTTPS://GITHUB.COM/GILLESW/CAPSTONE_IBM/BLOB/MASTER/MODULA%204%20ET%205/MODULE4-INTRO.IPYNB