Where to go if you visit ISSY-LES-MOULINEAUX city?

1. Introduction/Business Problem

1.1 Problem description:

Business men usually stay in non-attractive places: far from touristic areas, and usually in nearby towns. Reasons are related to business: customer/suppliers are usually located here, but also related to cost: Company proposes low cost hotels.

For diner and at night, where should he go?

Questions we need to answer here are:
- Is Foursquare enough for helping all travelers? (or Is foursquare only related to touristic places?)
- Where are the attractive venues in cities, when Foursquare is not enough?'

1.2 Data description:

There will be 2 datasets used for this study:
- Foursquare: List of POI (points of interest)
  Data are related to POI and each POI has a list of comments related to each of them.
- Open Data from Issy-Les-Moulineaux city hall: https://data.iledefrance.fr/explore/dataset/points-dinteret-de-la-ville-dissy-les-moulineaux/table/
  Data available here are linked to all activities related to this city: Companies (>50 employees), restaurants, public transport, ...

2. Data usage

2.1 Confirm Data Governance

Several years ago, the French government decided to grand access to their public datasets as 'open data'. Datas are free of charge, and contains many interesting items, especially related to companies and people. My planned capstone project was to combine Foursquare Data to Open Data from Paris city hall.

But while I tried to access data from the website https://data.iledefrance.fr , I read the 'confidentiality items' in detail and discovered, that some data are restricted: They are free, but you need to fill forms and apply procedures to declare usage... and this procedure takes long time and much efforts, which are not compatible with my Capstone Project.

Finally, I am proposing a capstone project in accordance with data governance's regulation.

2.2 Compare Data
The main difference between the two is:

- Foursquare contains advices from users, but might not be complete
- City's open-data includes all venues, but has no advices from customers

Questions are:
- Is Foursquare representative enough of the city? Do we have enough data?
- Can we combine databases? to have a complete view of existing venues, with advices?

Open Data

Here is an example of data from the city.

	Titre	Favoris	Catégorie 1	Catégorie 2	Catégorie 3	Catégorie 4	Description	Adresse	Codepostal	Ville	Téléphone	URL	Email	WGS84	Latitude	Longitude
0	Ben Yahia Birck	NaN	Vie pratique	Commerces, consommation	Epiceries, supérettes et grande distribution	NaN	Du mardi au dimanche de 8h à minuit	35, rue Ernest Renan	92130.0	Issy-les-Moulineaux	01 46 48 64 51	NaN	NaN	48.826655,2.280579	48.826655	2.280579
1	Conteneur à verre Boulevard Galliéni	NaN	Vie pratique	Environnement, cadre de vie	Conteneurs à verre	NaN	Conteneur enterré, à l'angle de la Rue Hoche.	83 Boulevard Galliéni	92130.0	Issy-les-Moulineaux	NaN	NaN	NaN	48.823029,2.267992	48.823029	2.267992
2	Autolib' boulevard Garibaldi	NaN	Vie pratique	Transports	Autolib'	NaN	Emplacement en voirie	18 boulevard Garibaldi	92130.0	Issy-les-Moulineaux	NaN	www.autolib.eu/fr/	NaN	48.821502,2.2613183	48.821550	2.261318
3	Parking motos Rue Diderot	NaN	Vie pratique	Transports	Parkings deux roues	Emplacements motos	Total de 4 places, côté pair à l'angle de la R...	Rue Diderot	92130.0	Issy-les-Moulineaux	NaN	NaN	NaN	48.825063,2.27213	48.825063	2.272130
4	Parking vélos Place Jacques Madaule	NaN	Vie pratique	Transports	Parkings deux roues	Emplacements vélos	Total de 12 places	Place Jacques Madaule	92130.0	Issy-les-Moulineaux	NaN	NaN	NaN	48.824084,2.260681	48.824084	2.260681

As we can see here, there is the venue name, location, but no data related to customers.
We do have the full list of venues (1188), but no advices from previous customers.

The categories are the following:

	Number of venues
Catégorie3	
Restauration-Bars-Tabacs	150
Parkings deux roues	125
Conteneurs à verre	63
Arbres remarquables	50
Bureaux de vote	46
Coiffeurs	32
Sculptures de métal	27
Epiceries, supérettes et grande distribution	27
Agences immobilières	26
Habillement	23

There are 150 bars and restaurants, which seems to be an attractive city… Now let's compare to Foursquare report.

Foursquare

Foursquare has a limitation, that this is the free version (and not the pro). Let's import ALL data around ISSY-LES-MOULINEAUX to see how much data we get.
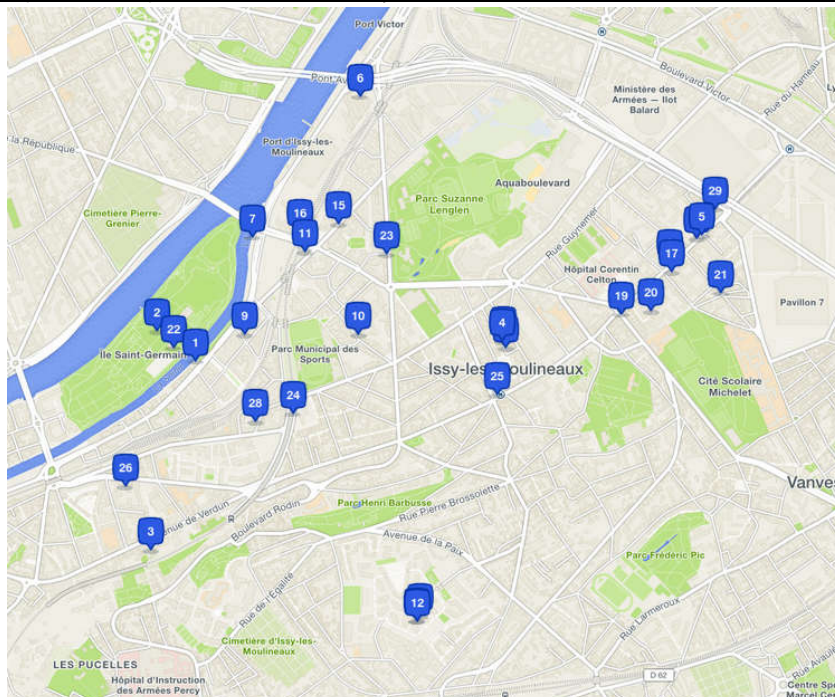
```
In [23]: Issy_venues.shape

Out[23]: (50, 5)
```

The API reported 50 venues only !
Lets confirm by going to the web site:

Foursquare seems poor of advices for this city.

Besides, a short comparison between the two databases shows, that GPS location and Venue title are not exactly the same.
Therefore, we cannot combine easily the 2 datasets.

2.3 AI

The city traveler wants to know where the best location is. To do this, he needs to know where people are going out for diner and at night.
By using data provided by the city itself, we can consider that restaurants and bars are located to the attractive places.
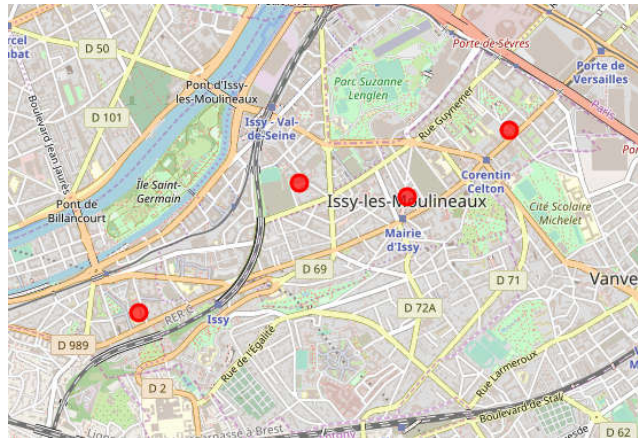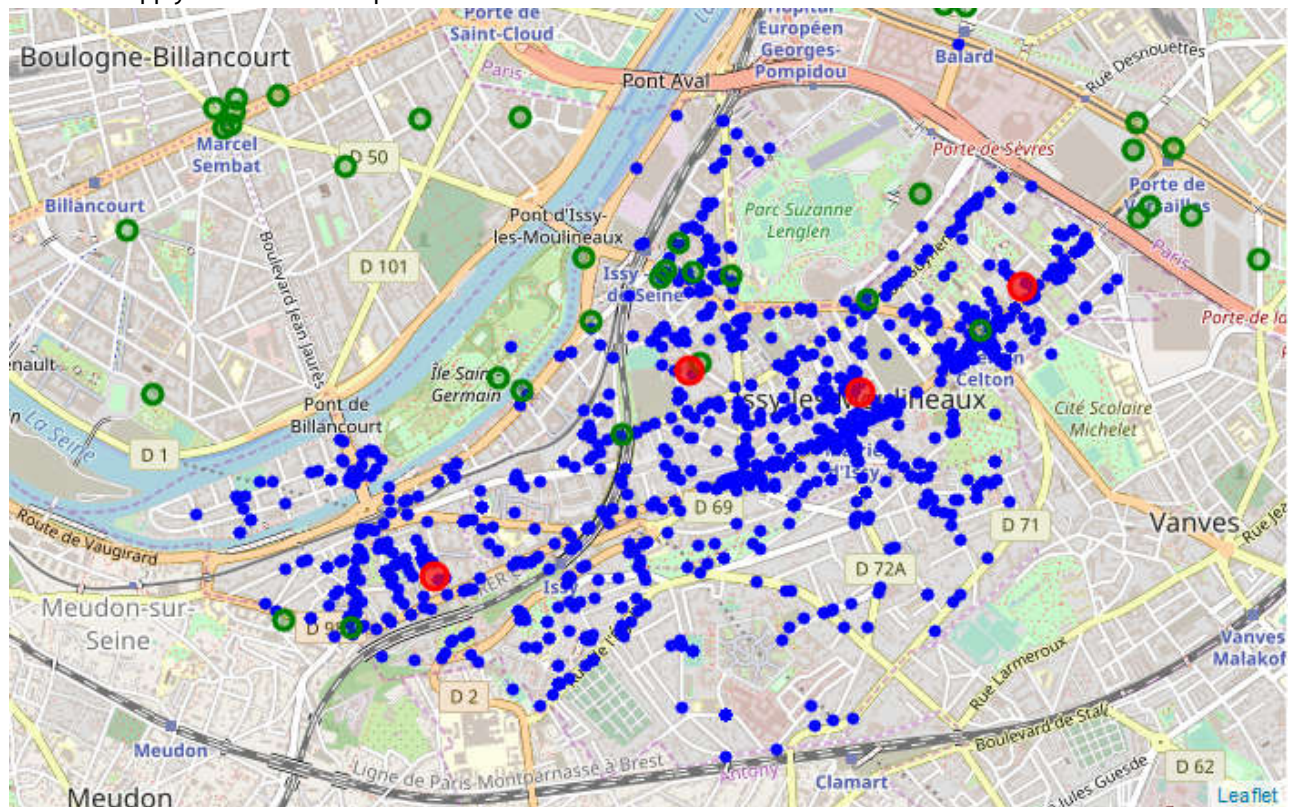Now let's consider only bars and restaurants…

To do this, we need to use a KNN algorithm: The Nearest Neighbor will give us the bests places.
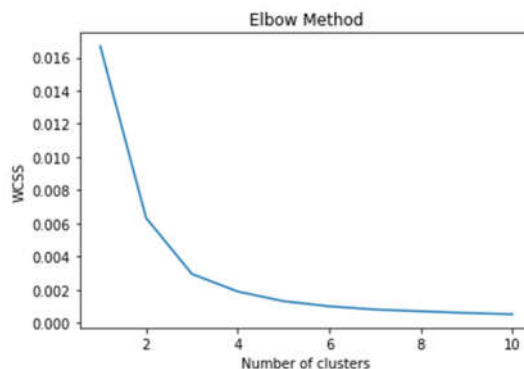
According to the Elbow method, only 4 locations are enough to describe the best places.

Now let's apply the found 'Best places' to the other venues from both datasets:

Green are the Foursquare data, Blue are the city open data, and Red are the KNN 'best places'.
(Open data considered here are the full base, not only restaurants)

Results for KNN application
By visual inspection, we confirm, that KNN really found the most attractive locations, whereas Foursquare proposes to take a taxi and drive to Boulogne-Billancourt ('Marcel Sembat' place).

3. Conclusion
In this Capstone Project, the approach proposes an overview on a global place.
By applying this method to other cities, we are able to compare cities to another...

For this particular city:
Foursquare has only few data, but does this mean this city is not interesting?
I believe, that French people are not used with FOURSQUARE. Only tourists do. And tourists focus on Paris…
and not on cities around it.
This would explain why Paris and other cities have many venues, and this Issy-Les-Moulineaux has not.
For information: maps.google.fr is well known and used… and therefore has much more data.

As a conclusion, I would say, that a travelers should use Foursquare for touristic places, but also should rely on other web sites to make its own opinion…