

WEBSCRAPING & APPLIED ML

*FIND YOUR
RESTAURANT*

MAEL DARNAUD
STELLA HU
HUGO LABORDE

SOMMAIRE

I. Webscraping

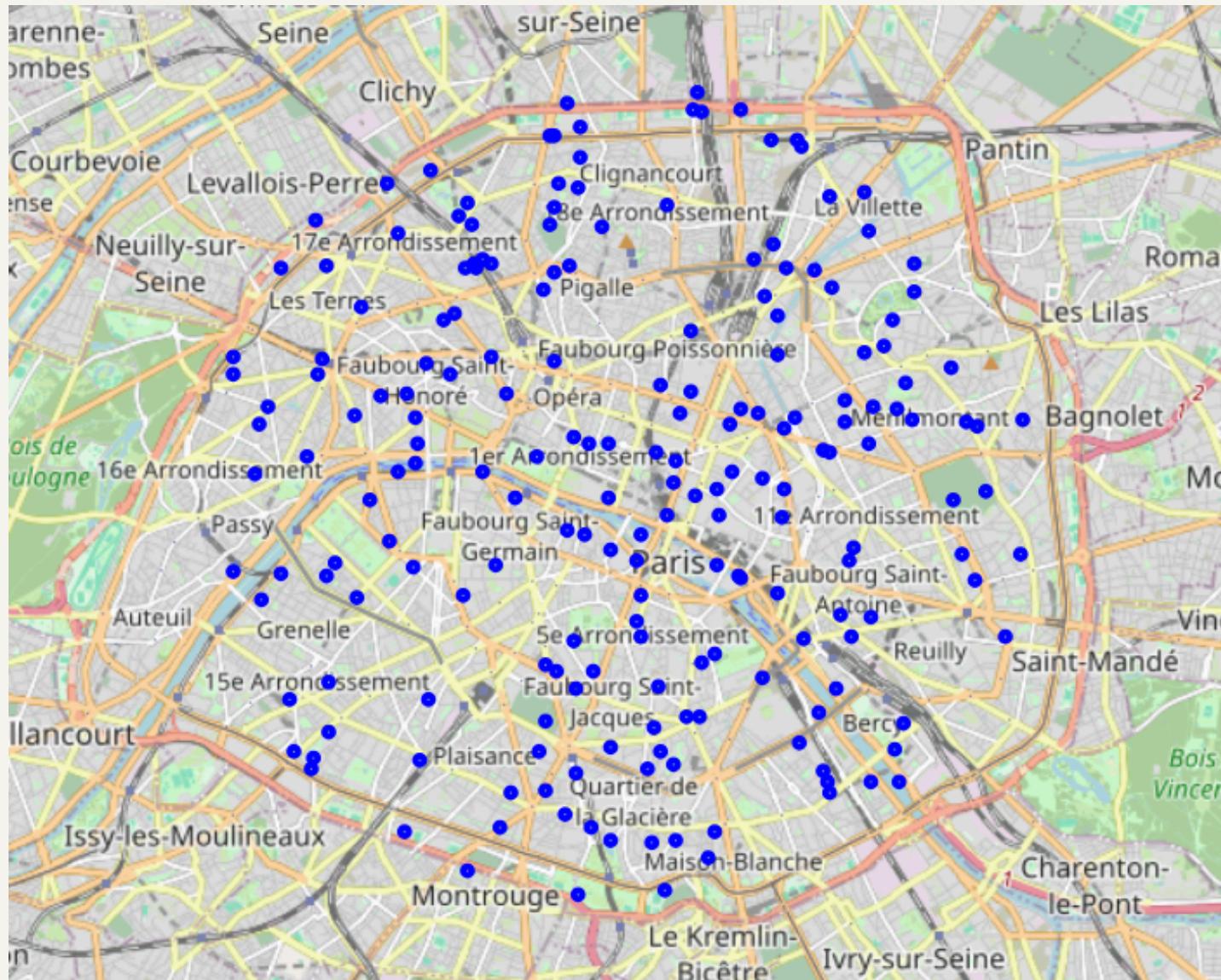
1. API

2. Selenium & bs4

II. Natural Language Processing

1. Restaurant recommendation system

WEBSRAPPIING WITH AN API: TRIPADVISOR



Generate random gps locations in
Paris

QUERY PARAMS

latLong string required
Latitude/Longitude pair to scope down the search around a specific point - eg. "42.3455,-71.10767"

key string required
The Partner API Key

category string
Filters result set based on property type. Valid options are "hotels", "attractions", "restaurants", and "geos"

phone string
Phone number to filter the search results by (this can be in any format with spaces and dashes but without the "+" sign at the beginning)

address string
Address to filter the search results by

radius string
Length of the radius from the provided latitude/longitude pair to filter results.

radiusUnit string
Unit for length of the radius. Valid options are "km", "mi", "m" (km=kilometers, mi=miles, m=meters)

language string Defaults to en
The language in which to return results (e.g. "en" for English or "es" for Spanish) from the list of our Supported Languages.

REQUEST

```
$ python -m pip install requests
1 import requests
2
3 url = "https://api.content.tripadvisor.com/api/v1/location/nearby"
4
5 headers = {"accept": "application/json"}
6
7 response = requests.get(url, headers=headers)
8
9 print(response.text)
```

Try It!

RESPONSE

```
1 {
2   "data": [
3     {
4       "location_id": "12453340",
5       "name": "Le Camondo",
6       "distance": "0.004612503290886536",
7       "bearing": "east",
8       "address_obj": {
9         "street1": "61bis Rue de Monceau",
10        "street2": "",
11        "city": "Paris",
12        "country": "France",
13        "postalcode": "75008",
14        "address_string": "61bis Rue de Monceau, 75008 Paris"
15      }
16    }
17  ]
```

200 LOG Headers ↗

Find the 10 nearest restaurant from a gps
location

WEBSRAPPIING WITH AN API

PATH PARAMS

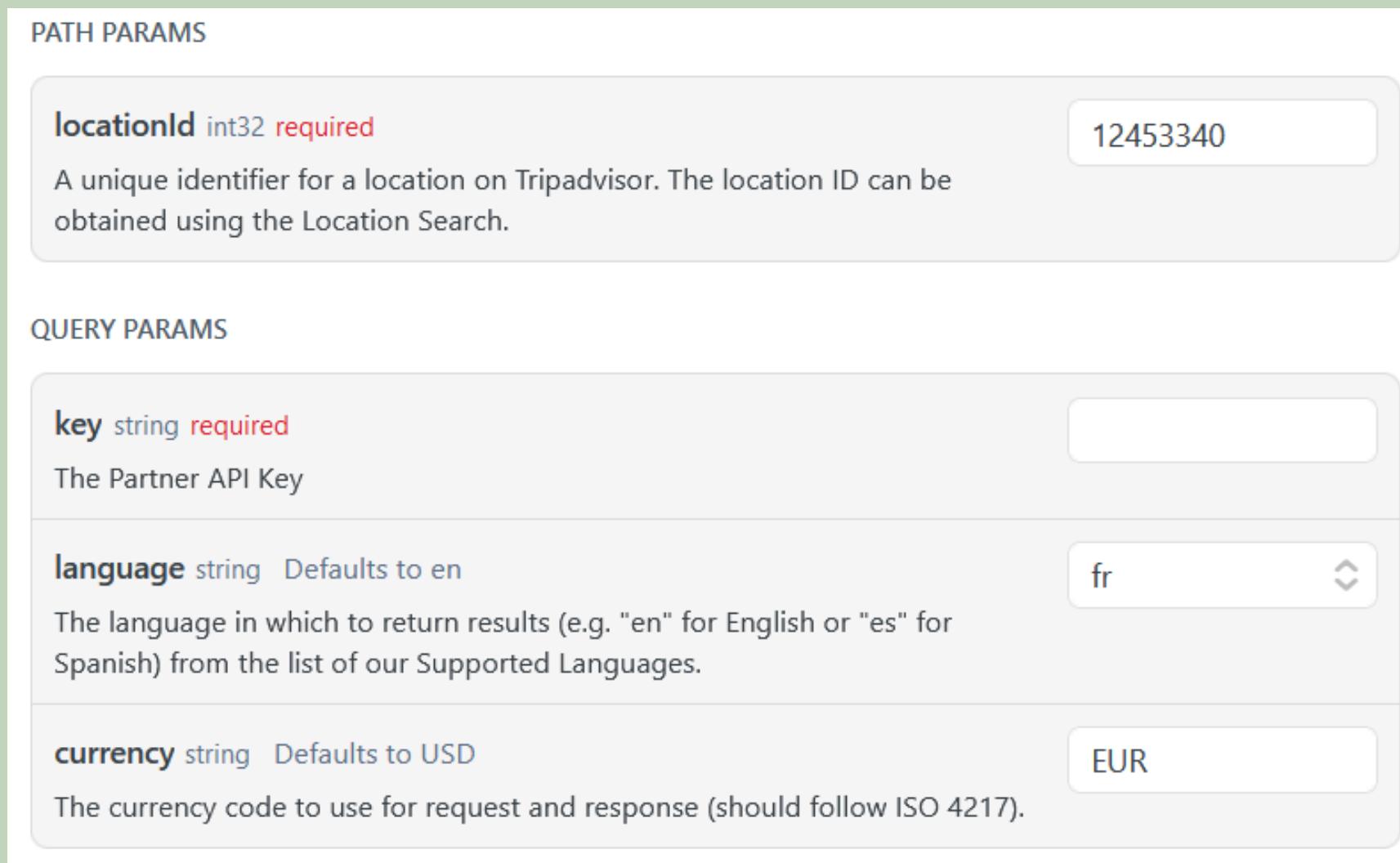
locationId int32 required
A unique identifier for a location on Tripadvisor. The location ID can be obtained using the Location Search.

QUERY PARAMS

key string required
The Partner API Key

language string Defaults to en
The language in which to return results (e.g. "en" for English or "es" for Spanish) from the list of our Supported Languages.

currency string Defaults to USD
The currency code to use for request and response (should follow ISO 4217).



Find details/reviews about restaurants
with locationId

REQUEST

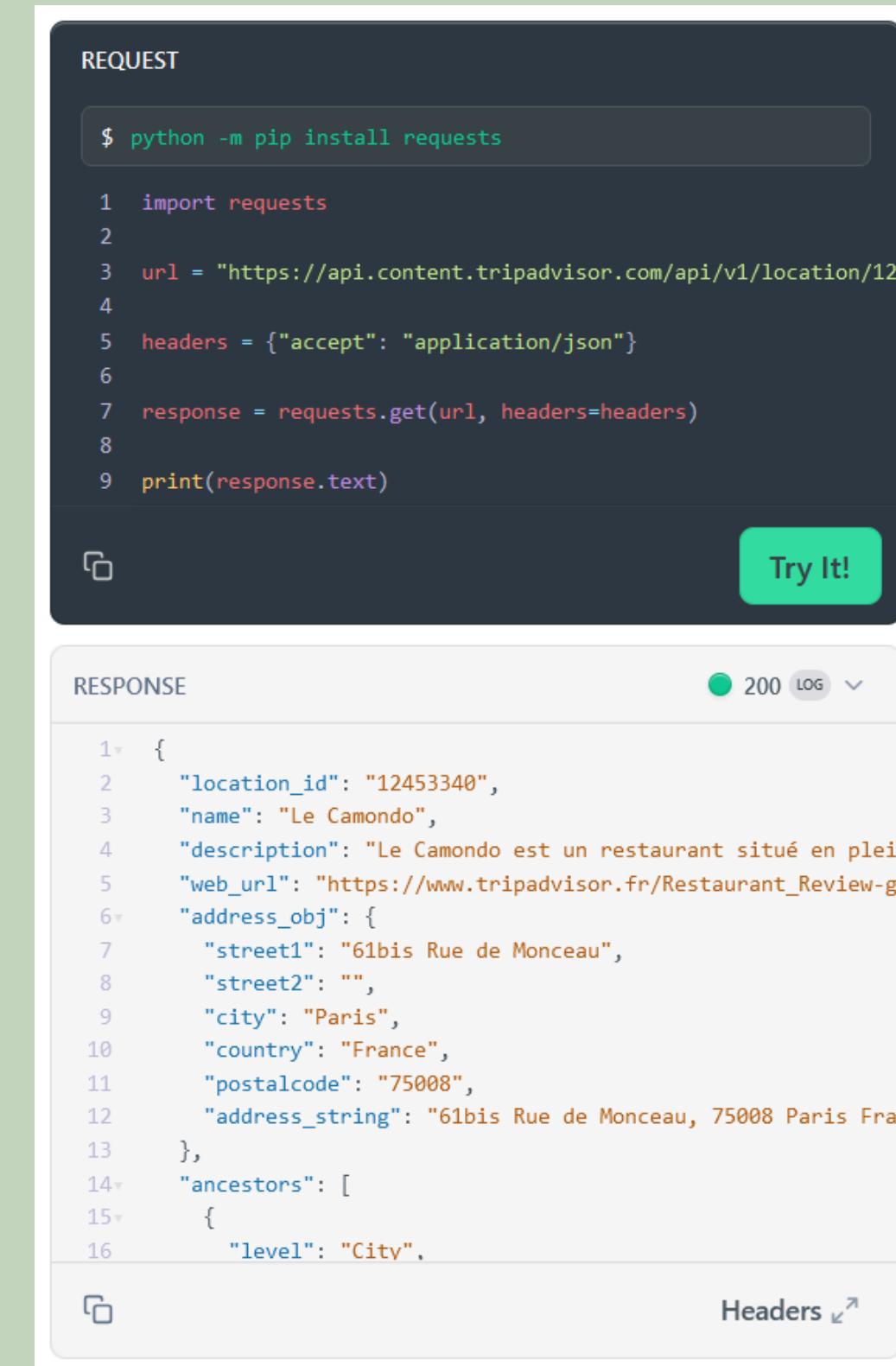
```
$ python -m pip install requests
1 import requests
2
3 url = "https://api.content.tripadvisor.com/api/v1/location/12453340"
4
5 headers = {"accept": "application/json"}
6
7 response = requests.get(url, headers=headers)
8
9 print(response.text)
```

Try It!

RESPONSE 200

```
1 {
2   "location_id": "12453340",
3   "name": "Le Camondo",
4   "description": "Le Camondo est un restaurant situé en plein cœur de Paris, offrant une cuisine française traditionnelle et contemporaine dans une atmosphère chaleureuse et conviviale. Les plats sont préparés à la commande avec des ingrédients frais et locaux. Le restaurant propose également une belle terrasse pour les beaux jours.",
5   "web_url": "https://www.tripadvisor.fr/Restaurant_Review-g187000-d12453340-Reviews-Le_Camondo-Paris.html",
6   "address_obj": {
7     "street1": "61bis Rue de Monceau",
8     "street2": "",
9     "city": "Paris",
10    "country": "France",
11    "postalcode": "75008",
12    "address_string": "61bis Rue de Monceau, 75008 Paris France"
13  },
14  "ancestors": [
15    {
16      "level": "City",
17      "name": "Paris"
18    }
19  ]
20}
```

Headers ↗



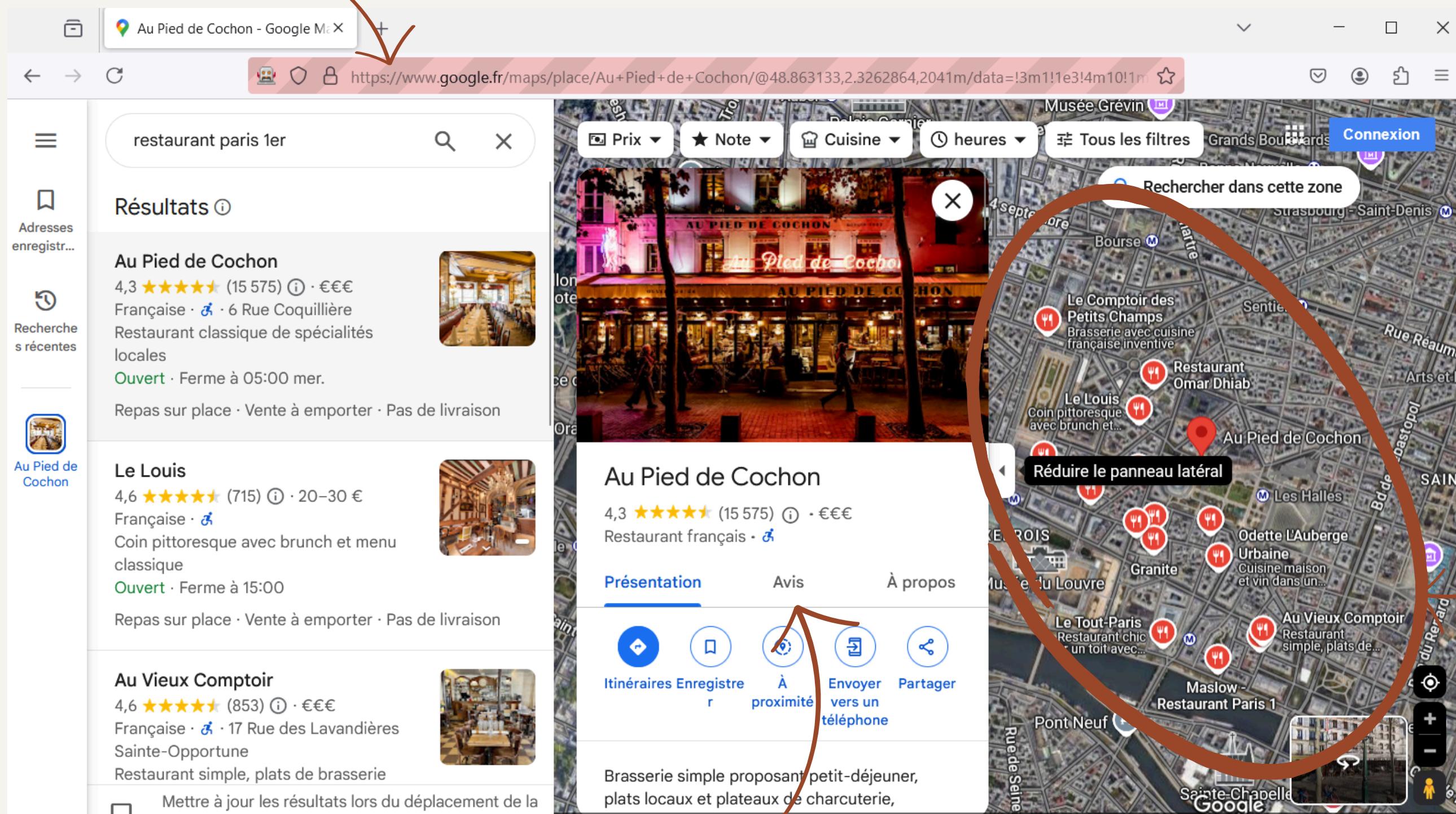
WEBSRAPPIING WITH AN API

	location_id	name	description	adress_string	price_level	rating	num_reviews	review_rating_count	cuisine
0	25940733	Ardent Paris	Table dédiée à la cuisine au feu avec un sourc...	40 Rue Richer, 75009 Paris France	€€ - €€€	4.5	24	{'1': '0', '2': '2', '3': '2', '4': '2', '5': ...}	Française
1	1116702	Casa Cosa	Situé en face du théâtre Mogador et à deux pas...	30, rue de Mogador, 75009 Paris France	€€ - €€€	4.0	475	{'1': '28', '2': '21', '3': '66', '4': '180', ...}	Italienne, Pizza, Méditerranéenne, Européenne,...
2	775346	Chez Vong	NaN	10 rue de la Grande Truanderie, 75001 Paris Fr...	€€€€	4.0	402	{'1': '30', '2': '54', '3': '51', '4': '112', ...}	Chinoise
3	15707789	Paye Ta Crêpe Paris 18e	NaN	92 Avenue de Saint Ouen, 75018 Paris France	NaN	3.5	6	{'1': '1', '2': '1', '3': '1', '4': '0', '5': ...}	NaN
4	8297103	McDonald's	Commandez en ligne et retrouvez facilement tou...	62 Boulevard Ornano, 75018 Paris France	€	3.0	4	{'1': '1', '2': '0', '3': '2', '4': '1', '5': ...}	NaN

Take only the data that interests us

WEBSRAPING WITH SELENIUM + BS4

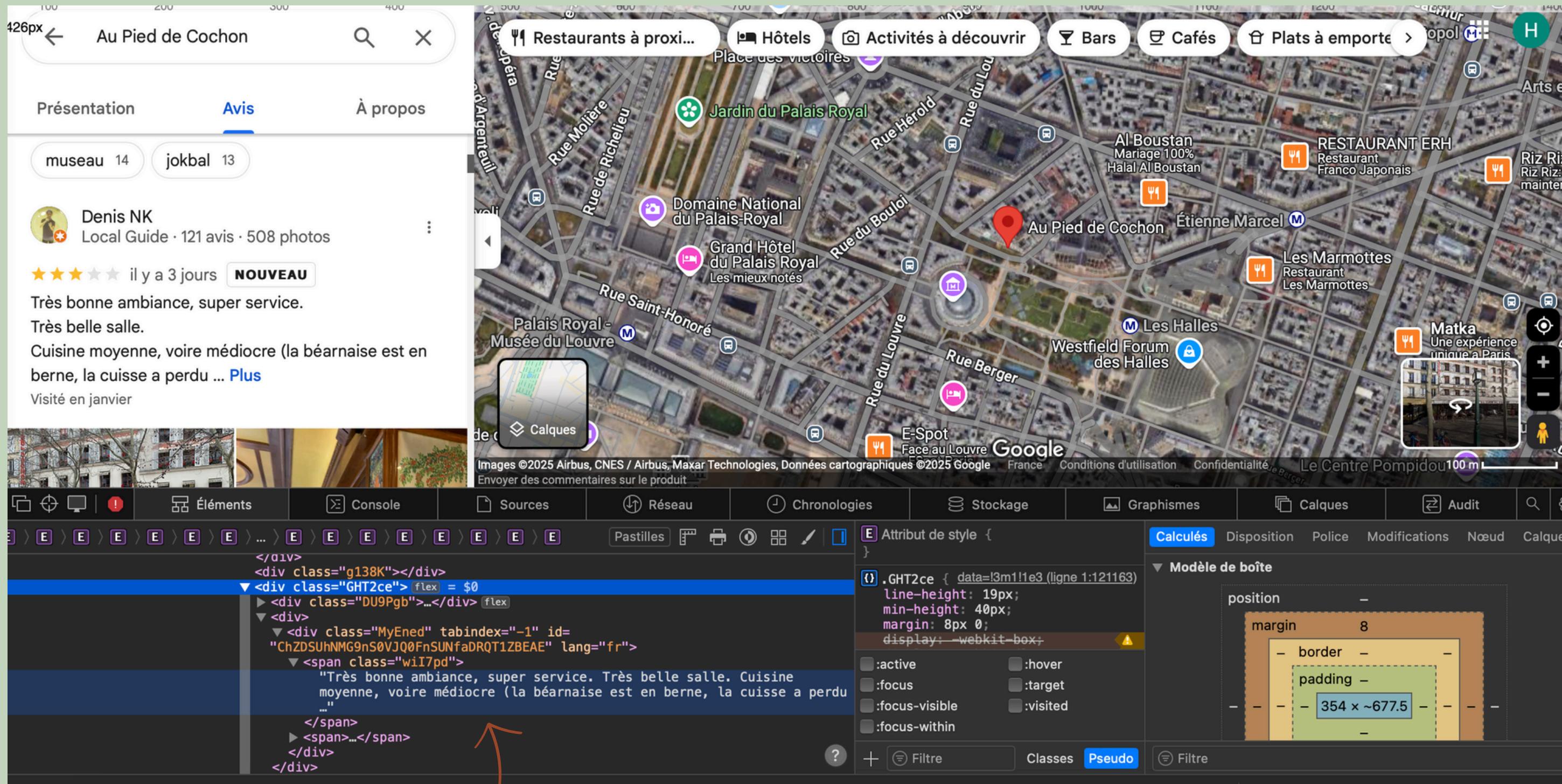
Step 2



Step 1

Step 3

WEBSCRAPPING WITH SELENIUM + BS4



Step 4

WEBSCRAPPING WITH SELENIUM + BS4

The image shows a screenshot of a web scraping process. At the top, there's a search bar with 'Au Pied de Cochon' and various filter buttons like 'Restaurants à proximité', 'Hôtels', 'Activités à découvrir', 'Bars', 'Cafés', and 'Plats à emporter'. Below the search bar is a map of a Parisian area with several restaurants marked. To the left of the map is a detailed view of a restaurant's page for 'Au Pied de Cochon'. The page includes tabs for 'Présentation', 'Avis' (which is selected), and 'À propos'. It shows reviews from users 'museau' (14) and 'jokbal' (13). A review by 'Denis NK' is highlighted, stating: 'Très bonne ambiance, super service. Très belle salle. Cuisine moyenne, voire médiocre (la béarnaise est en berne, la cuisse a perdu ... Plus'. Below the reviews is a photo of the restaurant's interior. On the right side of the interface, there's a code editor showing the HTML structure of the page, specifically focusing on a list of restaurants. At the bottom, there's a table with columns: name, rating, price, restaurant_type, description, address, and reviews. The table contains five rows of data.

	name	rating	price	restaurant_type	description	address	reviews
0	L'Aubergeade	4.8	€30–40	Restaurant	Typical, cozy bistro focusing on traditional F...	17 Rue Chaligny, 75012 Paris, France	["Un sans faute pour l'Aubergeade. Nous nous y...
1	La Cantine de daumesnil - restaurant paris 12	4.3	€20–30	Restaurant	NaN	119 Rue de Reuilly, 75012 Paris, France	['Je conseille pour les amoureux des plats ant...
2	Le Picotin	4.5	€20–40	French restaurant	Classic fare plated in a homey, relaxed settin...	35 Rue Sibuet, 75012 Paris, France	['Restaurant caché dans une ruelle calme.\nLes...
3	L'Alchimiste	4.7	€20–30	French restaurant	NaN	181 Rue de Charenton, 75012 Paris, France	['Cadre très agréable, accueil au top et repas...
4	Ty Louis - Restaurant Breton Paris 12	4.8	€20–40	Restaurant	NaN	240 Rue du Faubourg Saint-Antoine, 75012 Paris...	['Et non, ce n'est pas parce que c'est breton ...

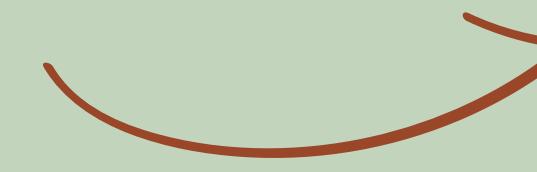
After processing with Python

FIND YOUR RESTAURANT: RECOMMENDATION SYSTEM

I. Data preprocessing

1. Text (reviews + description) cleaning
2. Translation in English
3. Lemmatization and stop word removing
4. Treatment to concatenate and make compatible the 2 datasets in one

description	reviews
Typical, cozy bistro focusing on traditional F...	["Un sans faute pour l'Aubergeade. Nous nous y...
NaN	['Je conseille pour les amoureux des plats ant...
Classic fare plated in a homey, relaxed settin...	['Restaurant caché dans une ruelle calme.\nLes...
NaN	["Cadre très agréable, accueil au top et repas...
NaN	['Et non, ce n'est pas parce que c'est breton ...



combined_desc_reviews
typical cozy bistro focusing traditional french...
advise lover west indian dish card well varied...
classic fare plated homey relaxed setting char...
pleasant framework top generous meal restauran...
breton pancake even à la carte pancake midi fo...

		name	rating	price	restaurant_type	description	address	reviews
0		L'Aubergeade	4.8	€30–40	Restaurant	Typical, cozy bistro focusing on traditional French...	17 Rue Chaligny, 75012 Paris, France	["Un sans faute pour l'Aubergeade. Nous nous y...
1		La Cantine de daumesnil - restaurant paris 12	4.3	€20–30	Restaurant	NaN	119 Rue de Reuilly, 75012 Paris, France	['Je conseille pour les amoureux des plats ant...
2		Le Picotin	4.5	€20–40	French restaurant	Classic fare plated in a homey, relaxed setting...	35 Rue Sibuet, 75012 Paris, France	['Restaurant caché dans une ruelle calme.\nLes...]
				€20	French		181 Rue de Charenton, 75012 Paris, France	["Cadre très agréable,



location_id	name	description	adress_string	price_level	rating	num_reviews	review_rating_count	cuisine
0	25940733 Ardent Paris	Table dédiée à la cuisine au feu avec un souci...	40 Rue Richer, 75009 Paris France	€€ - €€€	4.5	24	{"1": "0", "2": "2", "3": "2", "4": "2", "5": ...}	Française
1	1116702 Casa Cosa	Situé en face du théâtre Mogador et à deux pas...	30, rue de Mogador, 75009 Paris France	€€ - €€€	4.0	475	{"1": "28", "2": "21", "3": "66", "4": "180", ...}	Italienne, Pizza, Méditerranéenne, Européenne,...
2	775346 Chez Vong	Nan	10 rue de la Grande Truanderie, 75001 Paris Fr...	€€€€	4.0	402	{"1": "30", "2": "54", "3": "51", "4": "112", ...}	Chinoise
3	15707789 Paye Ta Crêpe Paris 18e	Nan	92 Avenue de Saint Ouen, 75018 Paris France	Nan	3.5	6	{"1": "1", "2": "1", "3": "1", "4": "0", "5": ...}	Nan

	id	text	title	rating	trip_type	location_id
0	987124639	I actually came every single time I've been in...	Asking for a tip	4	Solo travel	17227776
1	972571018	My daughter and i came here for a late lunch o...	Relaxing Lunch, great spot !	5	Family	17227776
2	948361293	Had a chocolate mousse with a couple of drinks...	Perfect cocktail spot	5	Couples	17227776
3	924882190	Loved this resraurant and it's location, food ...	Delicious French food	5	Couples	17227776
4	903705071	We had drinks several days in a row. Friendly ...	Nice place for drinks.	5	Family	17227776



		name	rating	price	restaurant_type	address	combined_desc_reviews_type
0		L'Aubergeade	4.8	40.0	Restaurant	17 Rue Chaligny, 75012 Paris 12	typical cozy bistro focusing traditional french...
1		La Cantine de daumesnil - restaurant paris 12	4.3	30.0	Restaurant	119 Rue de Reuilly, 75012 Paris 12	advise lover west indian dish card well varied...
2		Le Picotin	4.5	40.0	French restaurant	35 Rue Sibuet, 75012 Paris 12	classic fare plated homey relaxed setting char...
3		L'Alchimiste	4.7	30.0	French restaurant	181 Rue de Charenton, 75012 Paris 12	pleasant framework top generous meal restauran...
4		Ty Louis - Restaurant Breton Paris 12	4.8	40.0	Restaurant	240 Rue du Faubourg Saint-Antoine, 75012 Paris 12	breton pancake even à la carte pancake midi fo...

FIND YOUR RESTAURANT: RECOMMENDATION SYSTEM

II. Application

1. Use a pre-trained model like SentenceTransformer's **all-MiniLM-L6-v2** to convert restaurant descriptions + reviews into embeddings
2. Compute cosine similarity between user input and embeddings to find the closest matches
3. Consider similarity, price, and rating in a weighted score to rank restaurants + filtering by arrondissement
4. Streamlit interface



FIND YOUR RESTAURANT: RECOMMENDATION SYSTEM

Restaurant Recommender

Find the perfect restaurant in Paris based on your preferences!

Describe the type of restaurant you're looking for:

asian cuisine to eat with family and friends

What is your max budget? (€)

35

-

+

Minimal rating (1-5):

3.50

1.00

5.00

Preferred arrondissement (1-20):

ex 12 ...

Find Restaurants

Streamlit application interface

Here are your top recommendations, bon appétit :

	Restaurant Name	Rating (1-5)	Price (€)	Type of Cuisine	Address
0	Lucky Paris 3 - 人间烟火	4.9000	20	Asian restaurant	47 Rue au Maire, 75003 Paris 3
1	Yansai 3 - Restaurant asiatique Paris - Guabao Taiwanais	4.8000	20	Hot pot restaurant	88 Rue des Gravilliers, 75003 Paris 3
2	PÂTES THAÏ Place d'italie	4.8000	20	Thai restaurant	6 Rue Primatice, 75013 Paris 13
3	Tricotin	3.9000	20	Chinese restaurant	15 Av. de Choisy, 75013 Paris 13
4	HAO HAO 好好	3.9000	20	Asian restaurant	23 Av. de Choisy, 75013 Paris 13

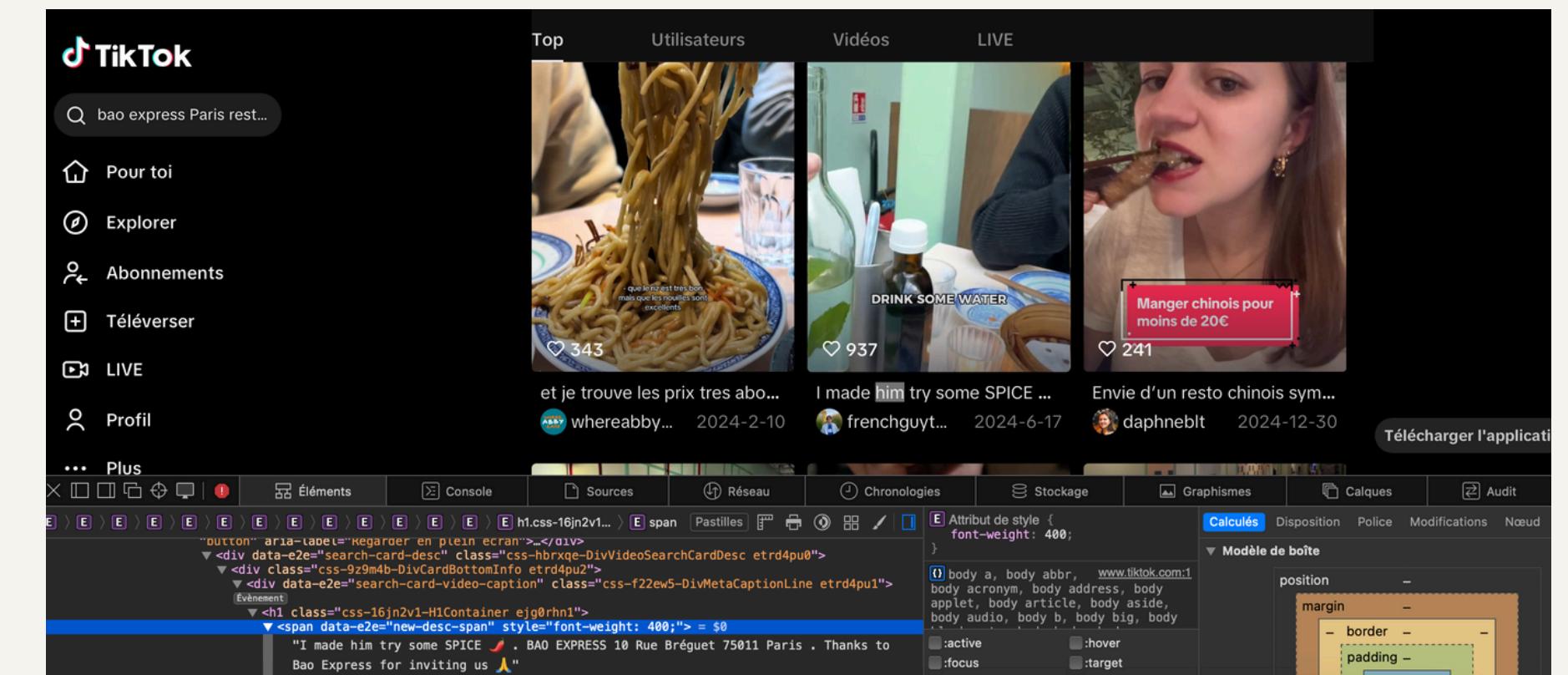
FUTURE PERSPECTIVES

- Use more ML systems
- One user query as a whole sentence
- More filters taken into account
- Graphic interface with more options (map)
- Quicker, less computational time
- Add a special filter “Seen on TikTok”

FUTURE PERSPECTIVES

- Add a special filter “Seen on TikTok”

1. With Selenium, find the search bar
2. Then type in the restaurant's name + Paris + restaurant
3. Scrap the description of the first video
4. With fuzzywuzzy, check if the description matches with the restaurant's name and address



Pas de pop-up de cookies.
Click search bar.
Recherche effectuée : Bao express paris restaurant
Descriptions TikTok récupérées :
1. I made him try some SPICE 🌶 . BAO EXPRESS 10 Rue Bréguet 75011 Paris . Thanks to Bao Express for inviting us 🙏

Name fuzz similarity score : 100
Address fuzz similarity score : 79
True

*THANK'S YOU
FOR
LISTENING*