

# Web Scrapping

Réalisé par :

- CANARIA Kaira
- FAREZ Lylia
- MAHOUCHE Tania
- ROLLET Gabrielle



<https://www.yelp.fr>



# Sommaire

1. Présentation du sujet
2. Chronologie des réalisations
3. Conception et démarches
4. Démo
5. Axes d'amélioration





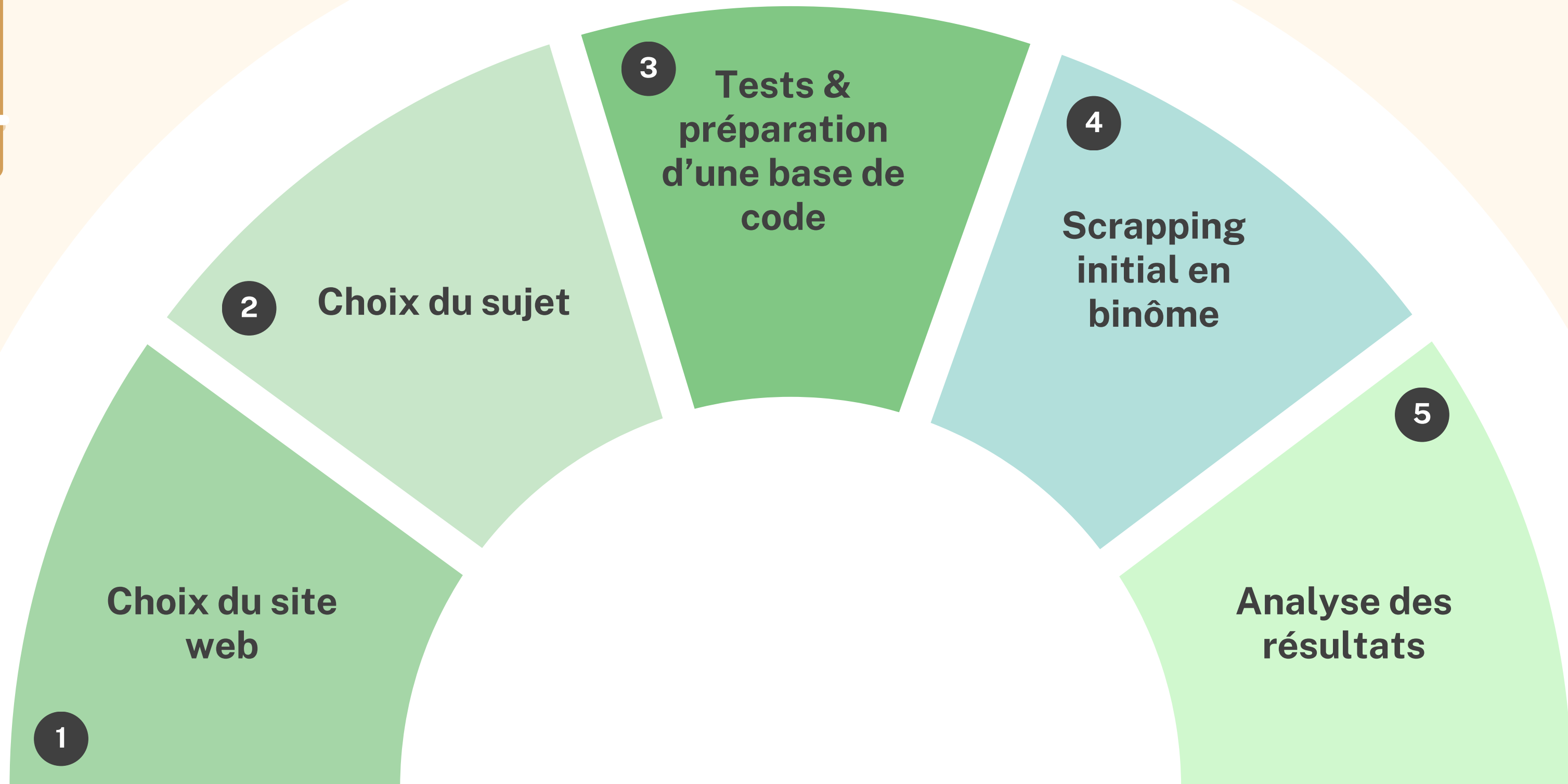
# Présentation du sujet



Ce projet consiste à scraper le site Yelp pour extraire les avis des utilisateurs et les utiliser comme base d'évaluation des restaurants afin d'analyser la satisfaction client et identifier les axes d'amélioration.



# Chronologie des réalisations



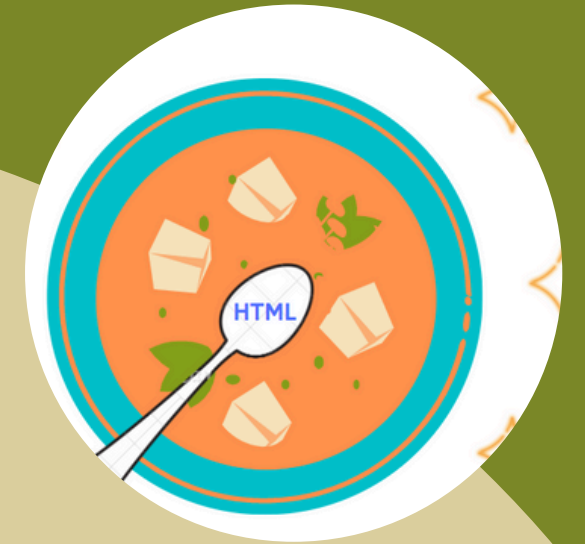
# Conception & démarches



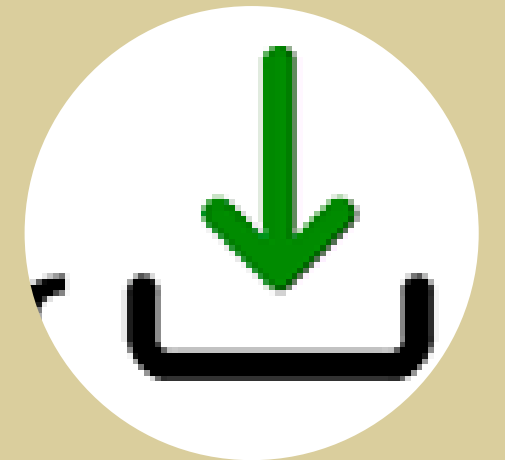
Selenium

## 1-Scrapping des données

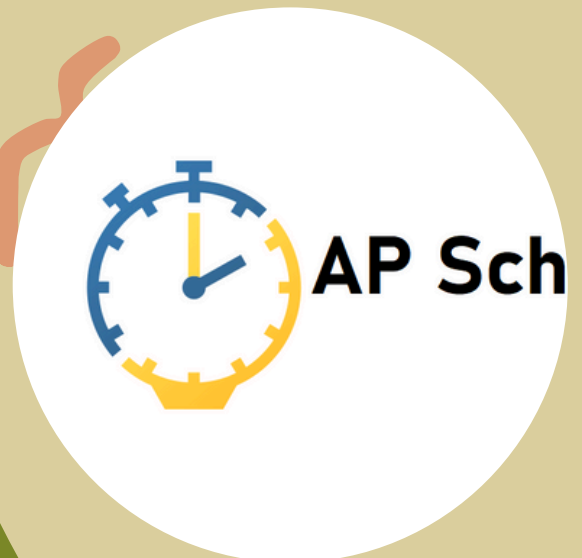
- Scrapping des restaurants
- Scrapping des commentaires



BeautifulSoup



Web driver  
manager



AP Scheduler

# Conception & démarches

## 2 – Stockage en base de données

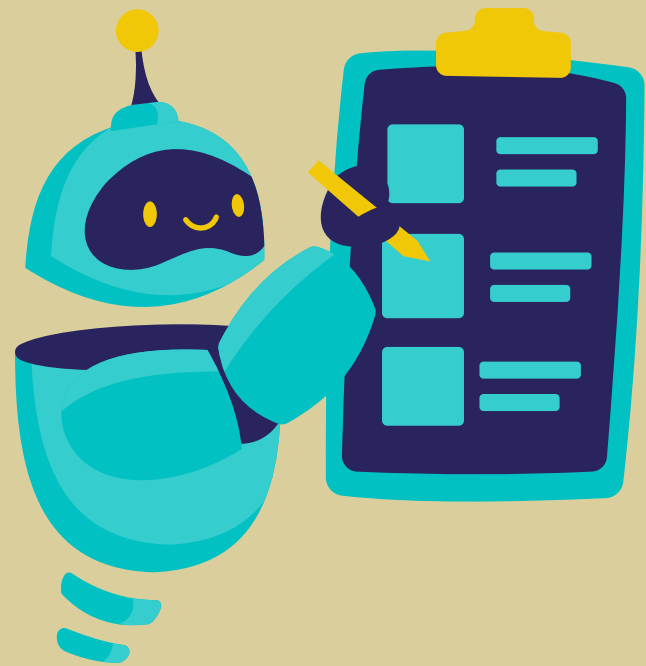
- Vérification de l'existence de la donnée

• Insertion



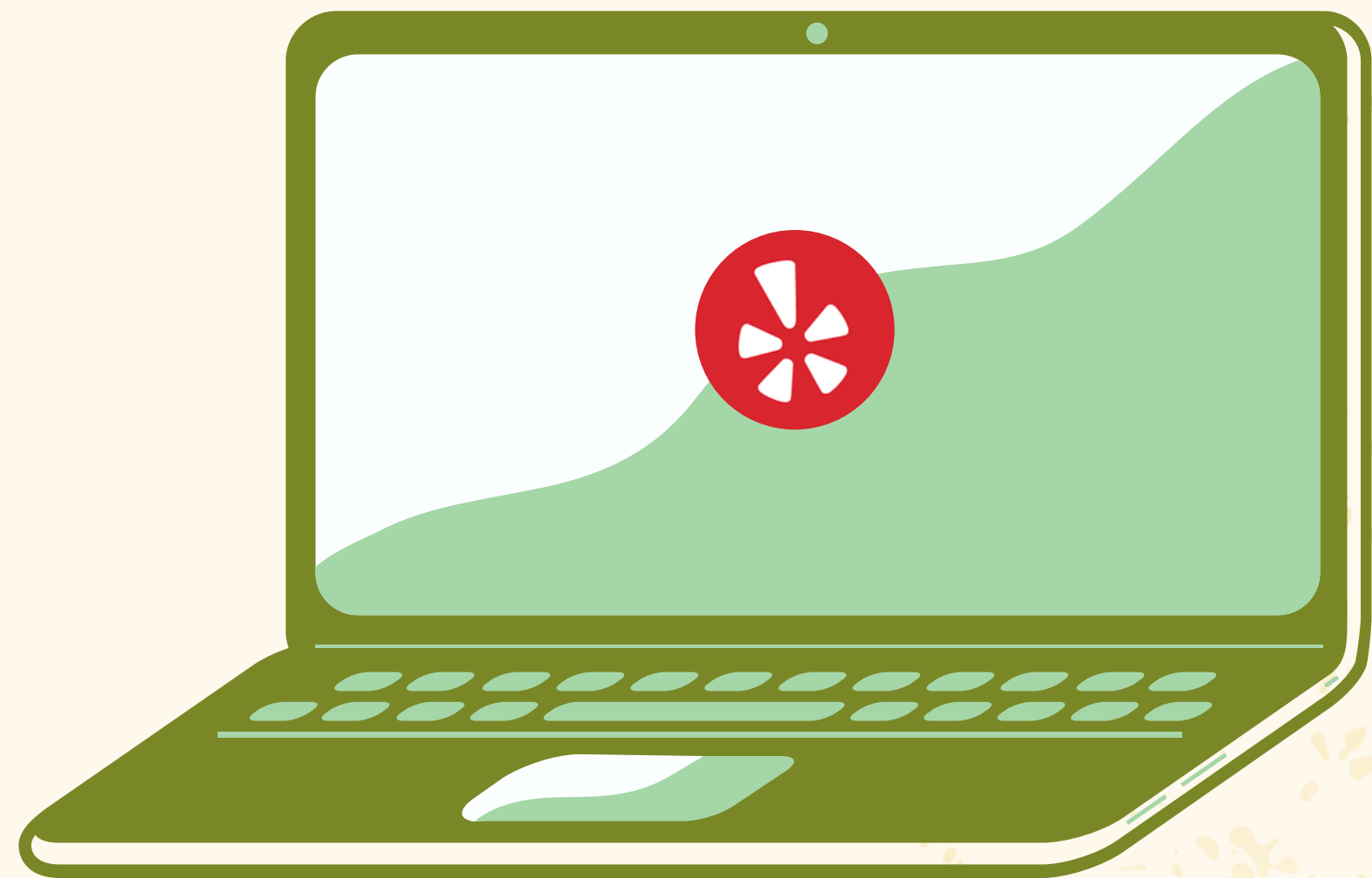
# Conception & démarches

## 3 – Prédiction & analyses



- Prédiction de la note moyenne (stars) d'un restaurant en fonction du nombre de commentaires (reviews) qu'il a reçus.
- Prédiction de la note des restaurants en fonction du sentiment moyen des commentaires .

Démo







# Axes d'amélioration

- Explorer de nouveaux modèles et prédire des éléments encore plus stratégiques pour les restaurants.
- Ajouter du threading pour le scrapping parallèle pour plus de rapidité et efficacité.
- Scrapping de tous les commentaires des restaurants.
- Ajout d'options de scrapping supplémentaire (choix du type d'établissement..)



**Merci pour votre attention**