

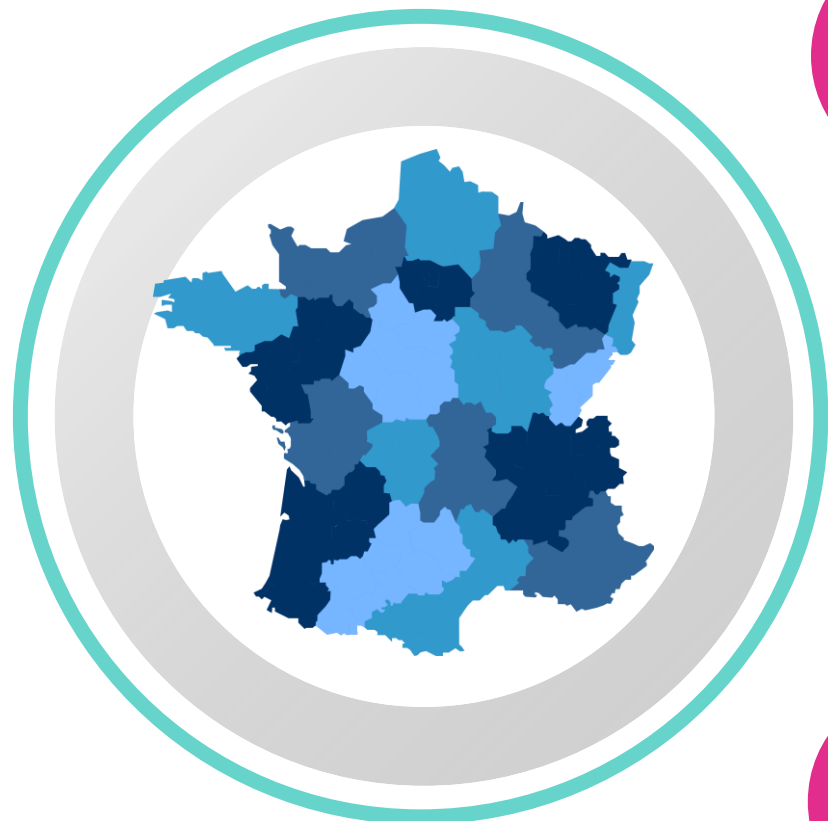


# ADN Tourisme

Projet 4 - Wild Code School

Sadjia, Lucile, Laurence, Zied, Victor

# Sommaire



1

Contexte & Objectifs

2

Méthodologie & Outils

3

Nos Conclusions

# 1 Le Contexte



ADN Tourisme est née le 11 mars 2020 du **regroupement des trois fédérations historiques des acteurs institutionnels du tourisme** : Offices de Tourisme de France, Tourisme & Territoires et Destination Régions



ADN Tourisme représente, au niveau national, les forces conjuguées de **1200 structures** et **13 500 salariés**.

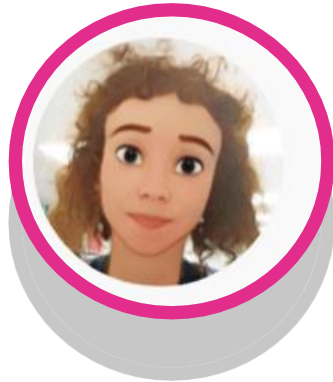
# L'Équipe

Sadjia Khadir



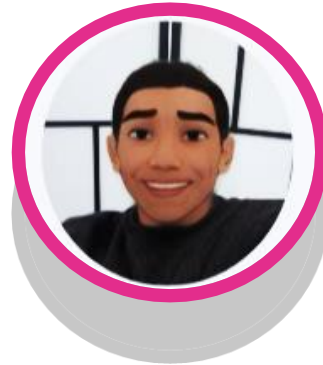
App  
Streamlit  
Manager

Lucile Lepage



Dashboard  
+ Shape  
Map Expert

Victor Estrada



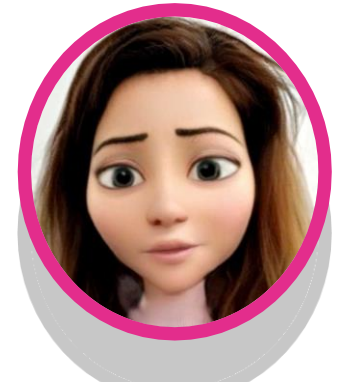
Machine  
Learning  
Manager

Zied Jouini



Flatten  
Wizard

Laurence Léon



Stagiaire  
Coerce

# 1 Les Délivrables et Objectifs



Un Prototype d'interface web



Comparer son offre

Un tableau de bord géographique permettant d'accéder aux informations touristiques



Montrer les tendances touristiques

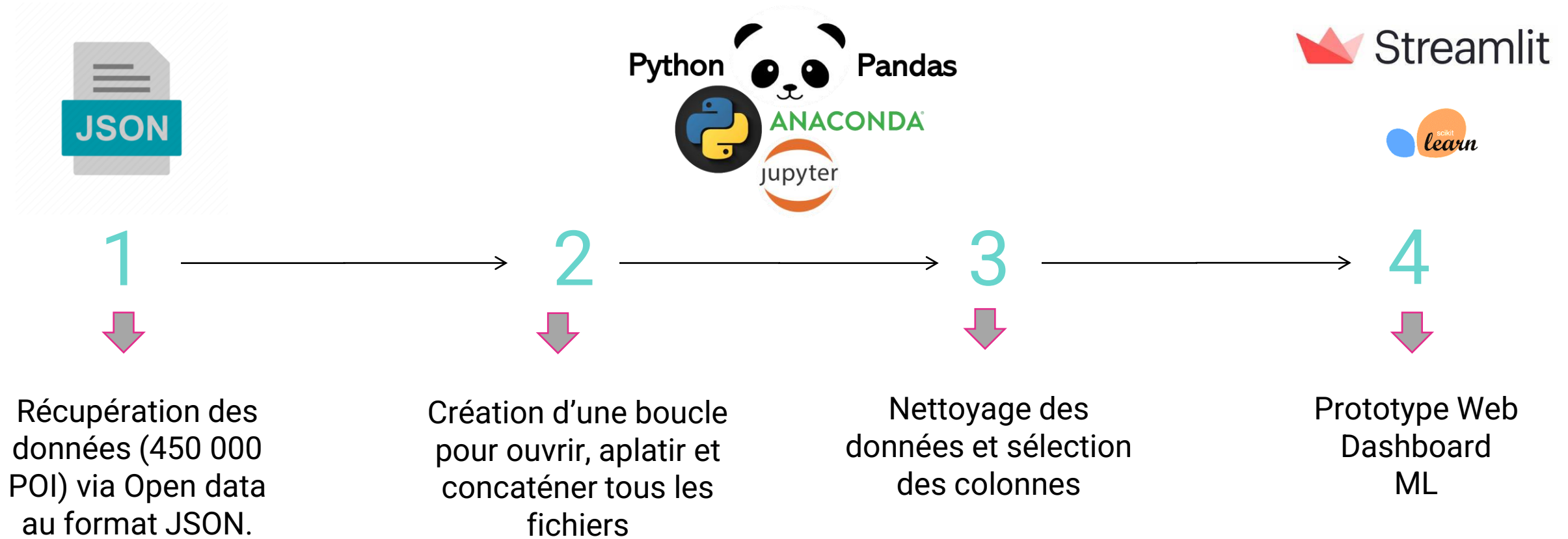
Prédire la catégorie en fonction de la description de l'établissement



Harmonisation automatique de la catégorie



## 2 La Méthodologie




## 2 Machine Learning: Catégorie\_Predictor

- L'intention de l'analyse ML est de prédire une catégorie de points d'intérêt
- 4 modèles de classification supervisée sont évalués
- Analyse basée sur les fonctionnalités NLP
- Model retenu : Logistic Regression


Model	Train/test score	Accuracy/R <sup>2</sup>
<b>Logistic Regression</b>	96.2/93.3	93.3
KNN Classification	93.9/91.4	91.4
Decision Tree Classifier	99.9/91.1	91.1
Random Forest Classifier	99.9/93.9	94.0

# Demo – POI Comparator

Choose your region  
Ile-de-France

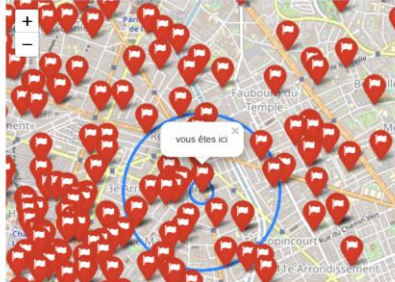
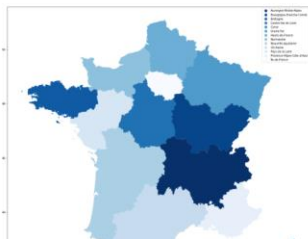

Choose your department  
Paris


Select your interest point


Welcome to Appli\_ADN\_Tourisme

HOME
**POI\_Comparator**
Dashboard
Category\_Predictor

Select your Point Of Interest



HOME
**POI\_Comparator**
Dashboard
Category\_Predictor


Select your Point Of Interest

Choose your department  
Paris

Select your interest point  
Adagio Paris Opéra

Select your information  

POI\_label x address x  
departement\_la... x  
description\_fr x email x  
postcode x region\_label x  
telephone x siret\_num x



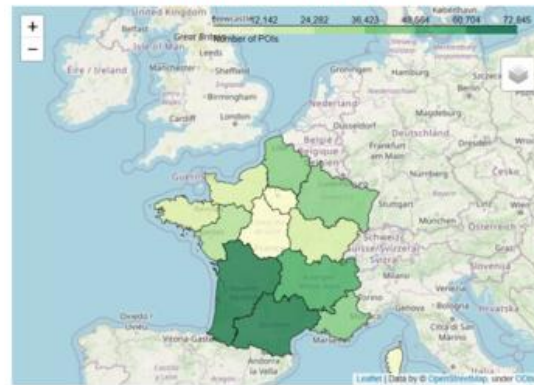
"POI\_Label" : "Adagio Paris Opéra"  
"description\_fr" :  
"Cette résidence de tourisme est idéalement installée en plein cœur de Paris, à quelques pas du Palais Garnier. Ce site exceptionnel, classé monument historique en 1923, met en scène ballets et concerts. A quelques minutes de marche, les Grands Magasins du boulevard Haussmann invitent au shopping. Les studios et appartements, décorés sobrement, sont tous dotés d'une cuisine. Une salle de sport est mise à disposition de la clientèle."  
"siret\_num" : "Information non renseignée"  
"email" : "H6988@adagio-city.com"  
"telephone" : "+33 1 58 18 08 20"  
"postcode" : "75009"  
"address" : "43 rue Caumartin"  
"departement\_label" : "Paris"  
"region\_label" : "Ile-de-France"



# Demo - Dashboard

[HOME](#)
[POI\\_Comparator](#)
[Dashboard](#)
[Category\\_Predictor](#)

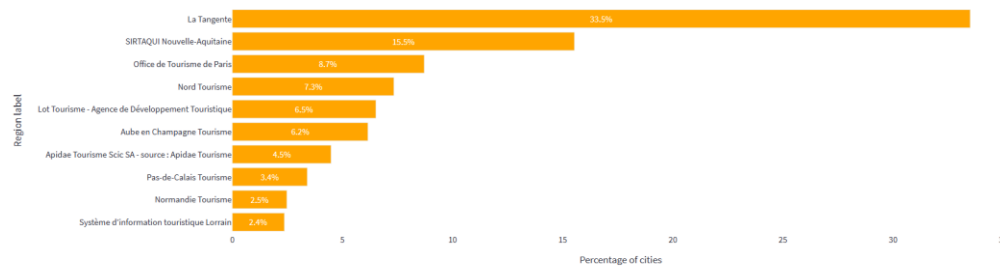
Points of interest by region



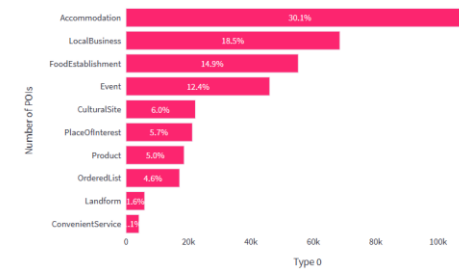
Points of interest by category



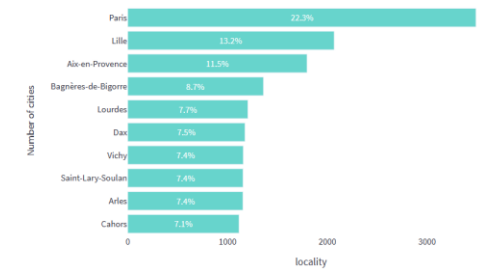
Absence of description by supplier



Top 10 by category



Top 10 by cities



# Demo – Catégorie\_Predictor

Machine learning time

The objective of this machine learning model is to verify the category and subcategories of a point of interest according to DATAtourism Ontology

The category is predicted using the establishment's name and description as input features for a supervised machine learning model (Logistic Regression)

This prototype has been deployed for accommodation and foodestablishment categories



Please, Enter the name and description of your establishment in French

Name:

haxo sushi

Description :

Click to see the category

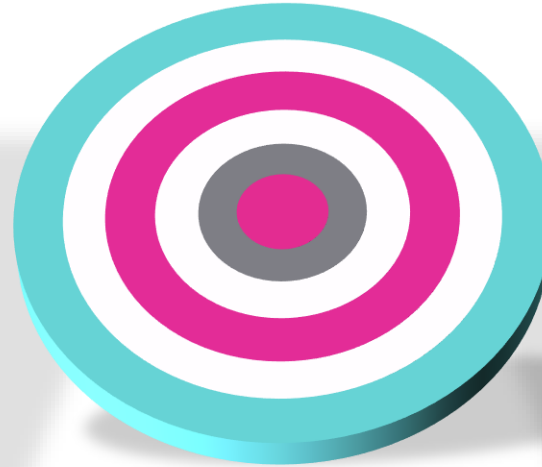
The main category of the POI is:

['FoodEstablishment']

# 3 Conclusion

## Nos Conclusions

- **30%** des POI appartiennent à la catégorie **hébergement**
- La Région la mieux référencée est la **Nouvelle Aquitaine**
- La Région la moins référencée est la **Guadeloupe**
- **34 086 POI (8%)** n'ont aucune description (parmi eux, la Tangente\*, agence des Hauts de France représente 33%)



## Nos Recommandations

- Harmoniser les champs à remplir dans chacune des offices de tourisme
- Collecter certaines données touristiques clefs (**labels et classifications**):
  - ✓ Prix
  - ✓ Accès Handicapés
  - ✓ Location Vélos
  - ✓ Villages fleuries/Villages de caractères...
- Systématiser l'utilisation du Machine Learning pour remplir les catégories