

# Projet 1 Openclassrooms

## Analyse des arbres de Paris



# Présentation du Projet

- **Contexte :**

Participation à un Challenge: dans le cadre du programme "**Végétons la ville**"

- **Objectif:**

Réalisez une analyse exploratoire sur le jeu de données sur les arbres de la ville de Paris

# Présentation du Projet

- **Règle du challenge :**

Installer un environnement virtuel dédié pour ce challenge

Les données doivent être exploitées avec python

- **Livrable attendu :**

Support de présentation de notre analyse

Jupyter notebook de l'analyse

# Environnement virtuel et choix des libraries

## 1. Environnement virtuel

- Création d'un environnement virtual en local sous ubuntu via Annaconda Nommé **Projet\_data\_arbre**
- Version python du projet : **python 3.11.4**

## 2. Bibliothèques utilisées :

- Pandas
- Seaborn
- Folium

# Présentation des données

- Jeu de données en Open data disponible sur le [site opendata paris](#). Données regroupant les arbres gérées par la ville de paris.
- Données disponible dans différents format via l'api ou téléchargement direct

# Contenu du Dataset

- Notre dataset contient les données de **207641** arbres de Paris

## Informations disponibles sur les arbres :

- ID
- Adresse
- Géolocalisation
- Espèces
- Arbre Remarquable
- Hauteur/circonférence
- Stade de développement

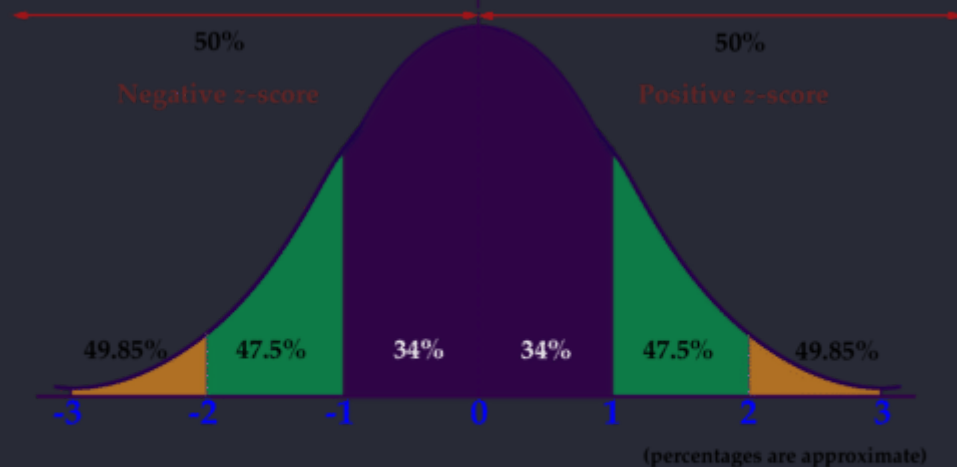
# Nettoyage des données : Suppression des colonnes inutiles

idbase	0.000000
typeemplacement	0.000000
domanialite	0.000000
arrondissement	0.000000
complementadresse	0.716068
numero	1.000000
adresse	0.000000
idemplacement	0.000000
libellefrancais	0.005803
genre	0.000010
espece	0.011934
varieteoucultivar	0.808723
circonferenceencm	0.000000
hauteuren	0.000000
stadedeveloppement	0.225399
remarquable	0.107060
geo_point_2d	0.000000
dtype: float64	

- Les colonnes **typeemplacement**, **numero**, **complementadresse** sont à supprimer

# Nettoyage des données : Suppression des valeurs abbérentes

## Le z-score





# Nettoyage des données : Suppression des valeurs abbérentes

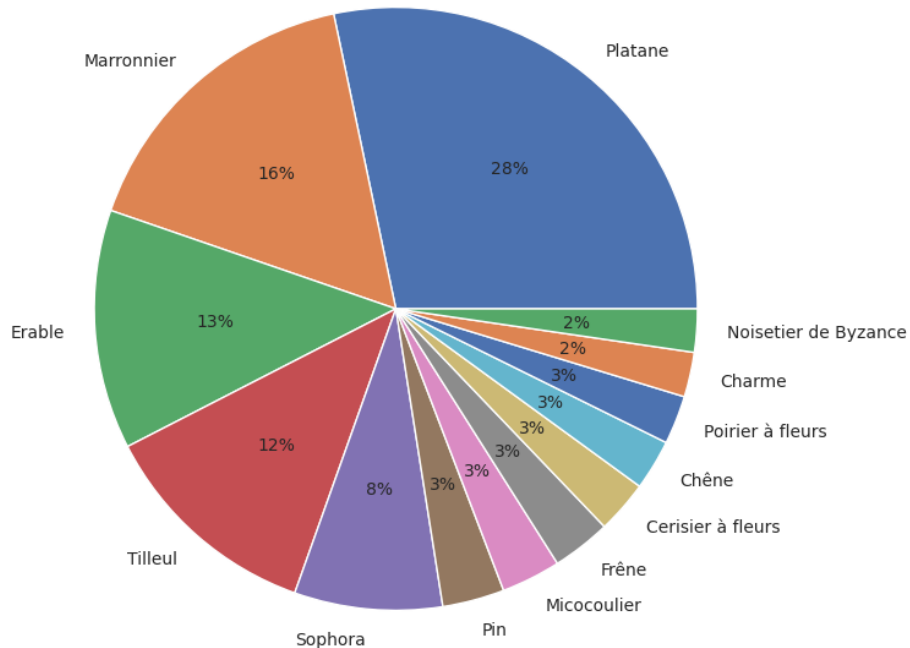
## Calcul du z-score sur notre jeu de données

- Z-score en fonction de la **hauteur** et la **Circoncéférence** des arbres
- Dataset contenant 2356 outliers

# Analyse de données

## Graphique: Représentation des Arbres dans paris

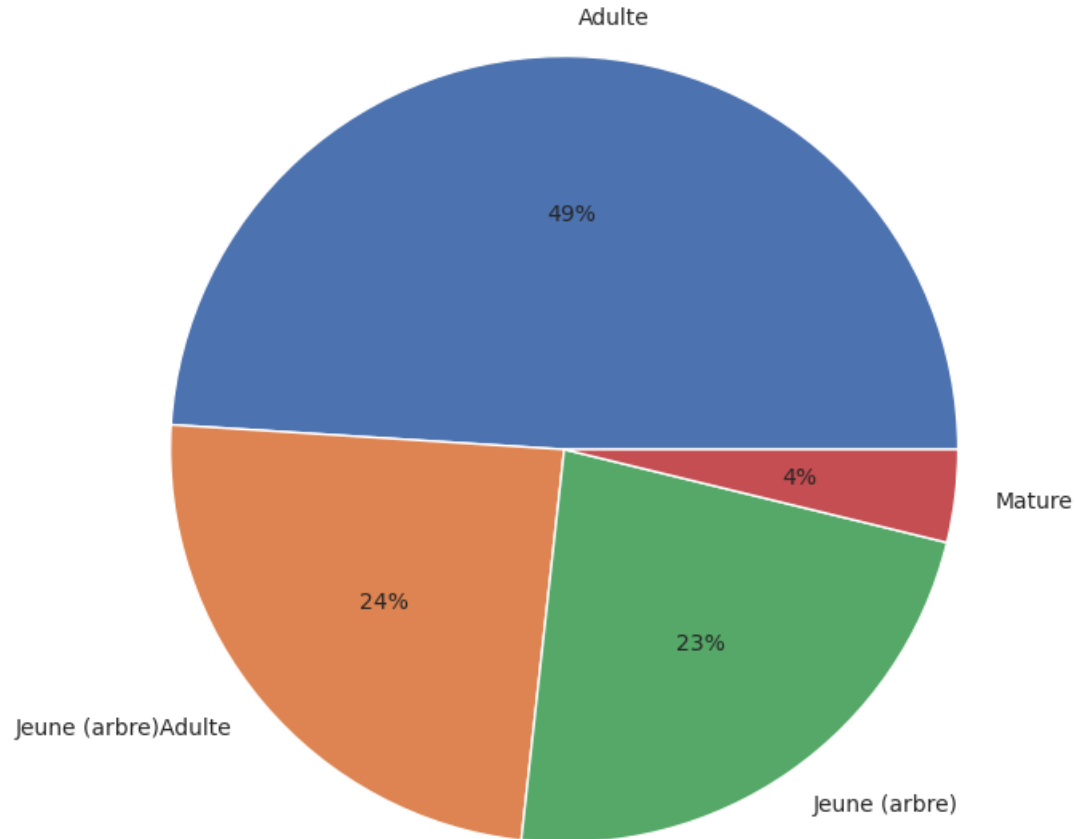
Réprésentation des espèces d'arbres à Paris > 3000 arbres



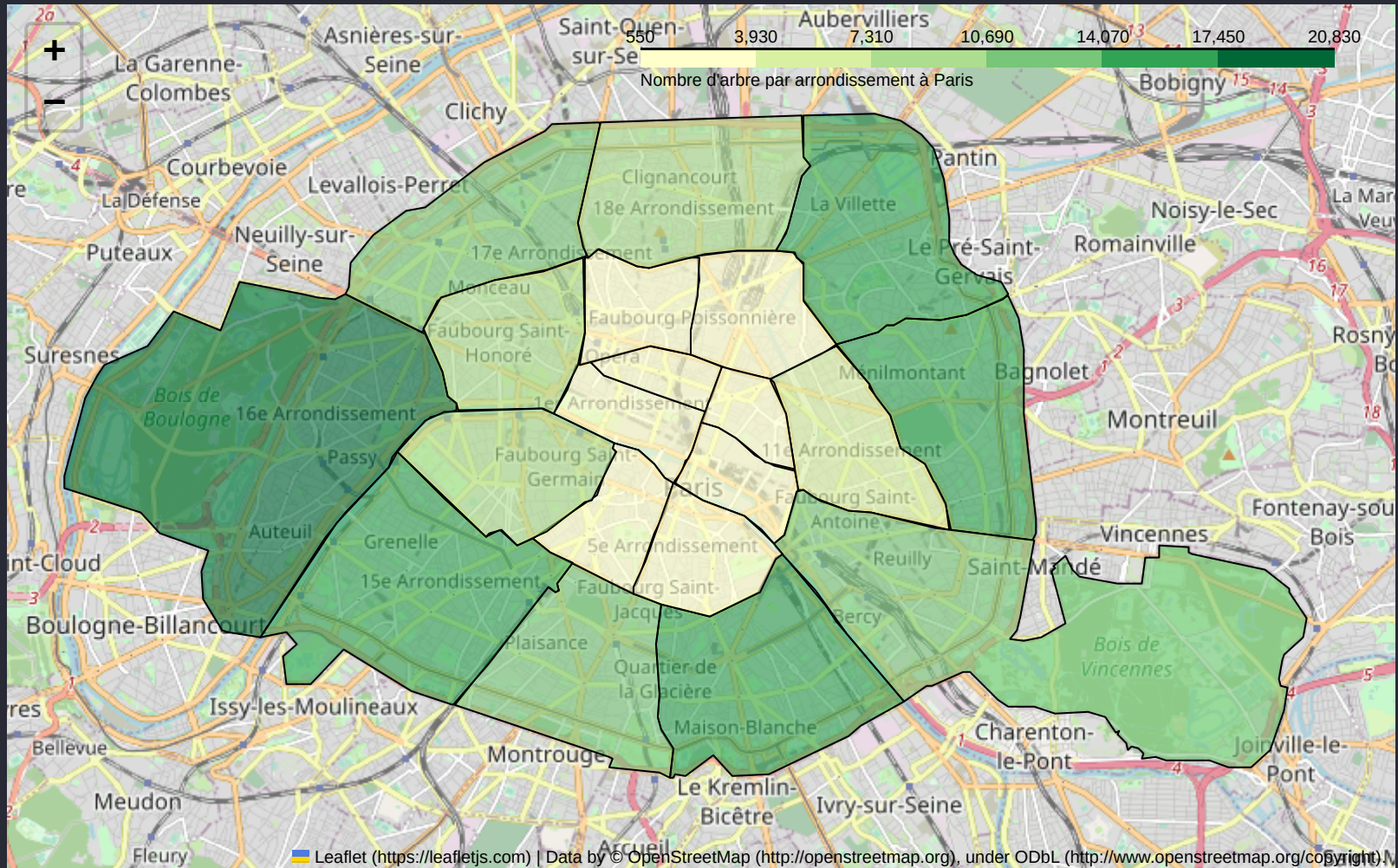
- **197 espèces** d'arbre différentes à paris
- Diversification des espèces : enjeux et climatique et de la survie des arbres

# Analyse de données Graphique: Représentation des Arbres dans paris

Réprésentation des arbres à Paris en fonction de leurs stade de développement



# Analyse Cartographique: Représentation des arbres par arrondissement



# Conclusion de l'analyse

