



Végétons la ville (Paris)

Sommaire

Contexte

Présentation générale du jeu de données

Analyse des données statistiques

Synthèse

Contexte



PARIS



OPTIMISATION
ENTRETIEN DES ARBRES



CHALLENGE SMART-
CITY

Création de l'environnement



Pour créer le fichier d'environnement :

```
conda env export > environment.yaml
```

Pour créer l'environnement :

```
conda env create --file environment.yaml
```



Présentation générale du jeu de données

Présentation du jeu de données brut



Dimensionnalité du DataFrame: (200137, 18)

	id	type_emplacement	domanialite	arrondissement	complement_adresse	numero	lieu	id_emplacement	libelle_francais	genre	espece	variete	circonference_cm	hauteur_m	stade_developpement	remarquable	geo_point_2d_a	geo_point_2d_b
0	99874	Arbre	Jardin	PARIS 7E ARRDT	NaN	NaN	MAIRIE DU 7E 116 RUE DE GRENELLE PARIS 7E	19	Marronnier	Aesculus	hippocastanum	NaN	20	5	NaN	0.0	48.857620	2.320962
1	99875	Arbre	Jardin	PARIS 7E ARRDT	NaN	NaN	MAIRIE DU 7E 116 RUE DE GRENELLE PARIS 7E	20	If	Taxus	baccata	NaN	65	8	A	NaN	48.857656	2.321031
2	99876	Arbre	Jardin	PARIS 7E ARRDT	NaN	NaN	MAIRIE DU 7E 116 RUE DE GRENELLE PARIS 7E	21	If	Taxus	baccata	NaN	90	10	A	NaN	48.857705	2.321061
3	99877	Arbre	Jardin	PARIS 7E ARRDT	NaN	NaN	MAIRIE DU 7E 116 RUE DE GRENELLE PARIS 7E	22	Erable	Acer	negundo	NaN	60	8	A	NaN	48.857722	2.321006
4	99878	Arbre	Jardin	PARIS 17E ARRDT	NaN	NaN	PARC CLICHY- BATIGNOLLES- MARTIN LUTHER KING	000G0037	Arbre à miel	Tetradium	daniellii	NaN	38	0	NaN	NaN	48.890435	2.315289

Contenu du fichier

```
RangeIndex: 200137 entries, 0 to 200136
Data columns (total 18 columns):
#   Column                Non-Null Count  Dtype  
---  -
0   id                     200137 non-null int64  
1   type_emplacement      200137 non-null object  
2   domanialite           200136 non-null object  
3   arrondissement        200137 non-null object  
4   complement_adresse    30902 non-null object  
5   numero               0 non-null     float64
6   lieu                 200137 non-null object  
7   id_emplacement        200137 non-null object  
8   libelle_francais      198640 non-null object  
9   genre                200121 non-null object  
10  espece                198385 non-null object  
11  variete               36777 non-null object  
12  circonference_cm       200137 non-null int64  
13  hauteur_m             200137 non-null int64  
14  stade_developpement   132932 non-null object  
15  remarquable           137039 non-null float64
16  geo_point_2d_a        200137 non-null float64
17  geo_point_2d_b        200137 non-null float64
dtypes: float64(4), int64(3), object(11)
memory usage: 27.5+ MB
```

Données numériques

	id	circonference_cm	hauteur_m	remarquable	geo_point_2d_a	geo_point_2d_b
count	200137.00000	200137.00000	200137.00000	137039.00000	200137.00000	200137.00000
mean	387202.68231	83.38048	13.11051	0.00134	48.85449	2.34821
std	545603.24236	673.19021	1971.21739	0.03662	0.03023	0.05122
min	99874.00000	0.00000	0.00000	0.00000	48.74229	2.21024
25%	155927.00000	30.00000	5.00000	0.00000	48.83502	2.30753
50%	221078.00000	70.00000	8.00000	0.00000	48.85416	2.35110
75%	274102.00000	115.00000	12.00000	0.00000	48.87645	2.38684
max	2024745.00000	250255.00000	881818.00000	1.00000	48.91148	2.46976

Données non numériques

	type_emplacement	domanialite	arrondissement	complement_adresse	lieu	id_emplacement	libelle_francais	genre	espece	variete	stade_developpement
count	200137	200136	200137	30902	200137	200137	198640	200121	198385	36777	132932
unique	1	9	25	3795	6921	69040	192	175	539	436	4
top	Arbre	Alignement	PARIS 15E ARRDT	SN°	PARC FLORAL DE PARIS / ROUTE DE LA PYRAMIDE	101001	Platane	Platanus	x hispanica	Baumannii'	A
freq	200137	104949	17151	557	2995	1324	42508	42591	36409	4538	64438

Analyse des données uniques

	unique_value	unique_count	row_not_na	unique_rate
id	[99874, 99875, 99876, 99877, 99878, 99879, 998...	200137	200137	100.0000
geo_point_2d_b	[2.3209621099, 2.3210314229, 2.321060566, 2.32...	200114	200137	99.9885
geo_point_2d_a	[48.8576199541, 48.8576560697, 48.8577050069, ...	200107	200137	99.9850
id_emplacement	[19, 20, 21, 22, 000G0037, 000G0036, 000G0035,...	69040	200137	34.4964
complement_adresse	[nan, c 12, 12-36, 12-35, face 64 Manin, 19-02...	3795	30902	12.2808
lieu	[MAIRIE DU 7E 116 RUE DE GRENELLE PARIS 7E, PA...	6921	200137	3.4581
variete	[nan, Atropunicea', Euchlora', Schwedleri', Pi...	436	36777	1.1855
espece	[hippocastanum, baccata, negundo, daniellii, x...	539	198385	0.2717
circonference_cm	[20, 65, 90, 60, 38, 37, 260, 145, 135, 15, 30...	531	200137	0.2653
libelle_francais	[Marronnier, If, Erable, Arbre à miel, Platane...	192	198640	0.0967
genre	[Aesculus, Taxus, Acer, Tetradium, Platanus, S...	175	200121	0.0874
hauteur_m	[5, 8, 10, 0, 17, 14, 3, 6, 4, 11, 7, 15, 9, 1...	143	200137	0.0715
arrondissement	[PARIS 7E ARRD, PARIS 17E ARRD, PARIS 16E AR...	25	200137	0.0125
domanialite	[Jardin, Alignement, DJS, DFPE, CIMETIERE, DAS...	9	200136	0.0045
stade_developpement	[nan, A, J, M, JA]	4	132932	0.0030
remarquable	[0.0, nan, 1.0]	2	137039	0.0015
type_emplacement	[Arbre]	1	200137	0.0005
numero	[nan]	0	0	NaN

Pourcentage des valeurs manquantes :

id	0.000000
hauteur_m	0.000000
circonference_cm	0.000000
geo_point_2d_a	0.000000
id_emplacement	0.000000
lieu	0.000000
geo_point_2d_b	0.000000
arrondissement	0.000000
type_emplacement	0.000000
domanialite	0.000500
genre	0.007995
libelle_francais	0.747988
espece	0.875400
remarquable	31.527404
stade_developpement	33.579498
variete	81.624088
complement_adresse	84.559577
numero	100.000000
dtype: float64	

Analyse des données manquantes

	genre	domanialite	espece	variete	genre	espece	circonference_cm	hauteur_m	stade_developpement
195404	NaN	Jardin	NaN	NaN	NaN	NaN	0	0	NaN
195409	NaN	Jardin	NaN	NaN	NaN	NaN	0	0	NaN
195410	NaN	Jardin	NaN	NaN	NaN	NaN	0	0	NaN
195475	NaN	Jardin	NaN	NaN	NaN	NaN	0	0	NaN
195476	NaN	Jardin	NaN	NaN	NaN	NaN	0	0	NaN
195487	NaN	Jardin	NaN	NaN	NaN	NaN	0	0	NaN
195496	NaN	Jardin	NaN	NaN	NaN	NaN	0	0	NaN
195497	NaN	Jardin	NaN	NaN	NaN	NaN	0	0	NaN
195499	NaN	Jardin	NaN	NaN	NaN	NaN	0	0	NaN
195502	NaN	Jardin	NaN	NaN	NaN	NaN	0	0	NaN
195503	NaN	Jardin	NaN	NaN	NaN	NaN	0	0	NaN
195505	NaN	Jardin	NaN	NaN	NaN	NaN	0	0	NaN
195509	NaN	Jardin	NaN	NaN	NaN	NaN	0	0	NaN
195514	NaN	Jardin	NaN	NaN	NaN	NaN	0	0	NaN
195518	NaN	Jardin	NaN	NaN	NaN	NaN	0	0	NaN
195519	NaN	Jardin	NaN	NaN	NaN	NaN	0	0	NaN

Analyse des données manquantes (suite)

Analyse des doublons

	duplicate_value_count	row_not_na	duplicate_rate
type_emplacement	200136	200137	100.000
remarquable	137037	137039	99.999
stade_developpement	132928	132932	99.997
domanialite	200127	200136	99.996
arrondissement	200112	200137	99.988
hauteur_m	199994	200137	99.929
genre	199946	200121	99.913
libelle_francais	198448	198640	99.903
circonference_cm	199606	200137	99.735
espece	197846	198385	99.728
variete	36341	36777	98.814
lieu	193216	200137	96.542
complement_adresse	27107	30902	87.719
id_emplacement	131097	200137	65.504
geo_point_2d_a	30	200137	0.015
geo_point_2d_b	23	200137	0.011
id	0	200137	0.000
numero	0	0	NaN

Analyse des doublons (suite)

	domanialite	id_emplacement	lieu	genre	variete	libelle_francais	geo_point_2d_a	geo_point_2d_b	stade_developpement
16	Jardin	190042	PARC FLORAL DE PARIS / ROUTE DE LA PYRAMIDE	Populus	NaN	Peuplier	48.836416	2.446277	NaN
17	Jardin	190042	PARC FLORAL DE PARIS / ROUTE DE LA PYRAMIDE	Populus	Italica'	Peuplier	48.836416	2.446277	A

	domanialite	id_emplacement	lieu	genre	variete	libelle_francais	geo_point_2d_a	geo_point_2d_b	stade_developpement
9	Alignement	402029	ROUTE DAUPHINE	Tilia	NaN	Tilleul	48.829912	2.439665	NaN
10	Alignement	402030	ROUTE DAUPHINE	Tilia	NaN	Tilleul	48.829912	2.439665	NaN

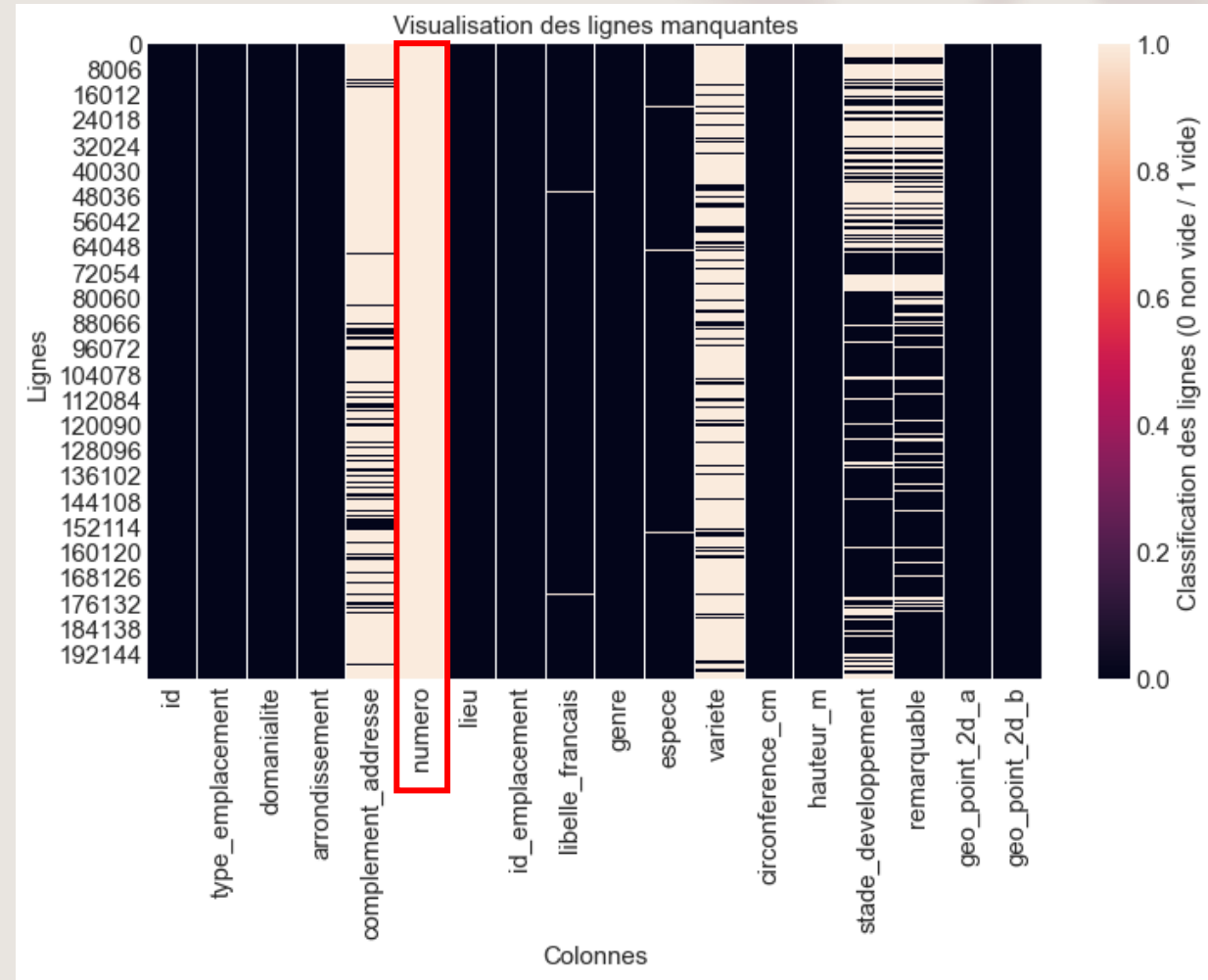
	domanialite	id_emplacement	lieu	genre	variete	libelle_francais	geo_point_2d_a	geo_point_2d_b	stade_developpement
5	Jardin	3	VOIE DF19	Populus	NaN	Peuplier	48.900253	2.386444	NaN
8	PERIPHERIQUE	09VO19003	VOIE DF/19	Populus	Italica'	Peuplier	48.900253	2.386444	NaN

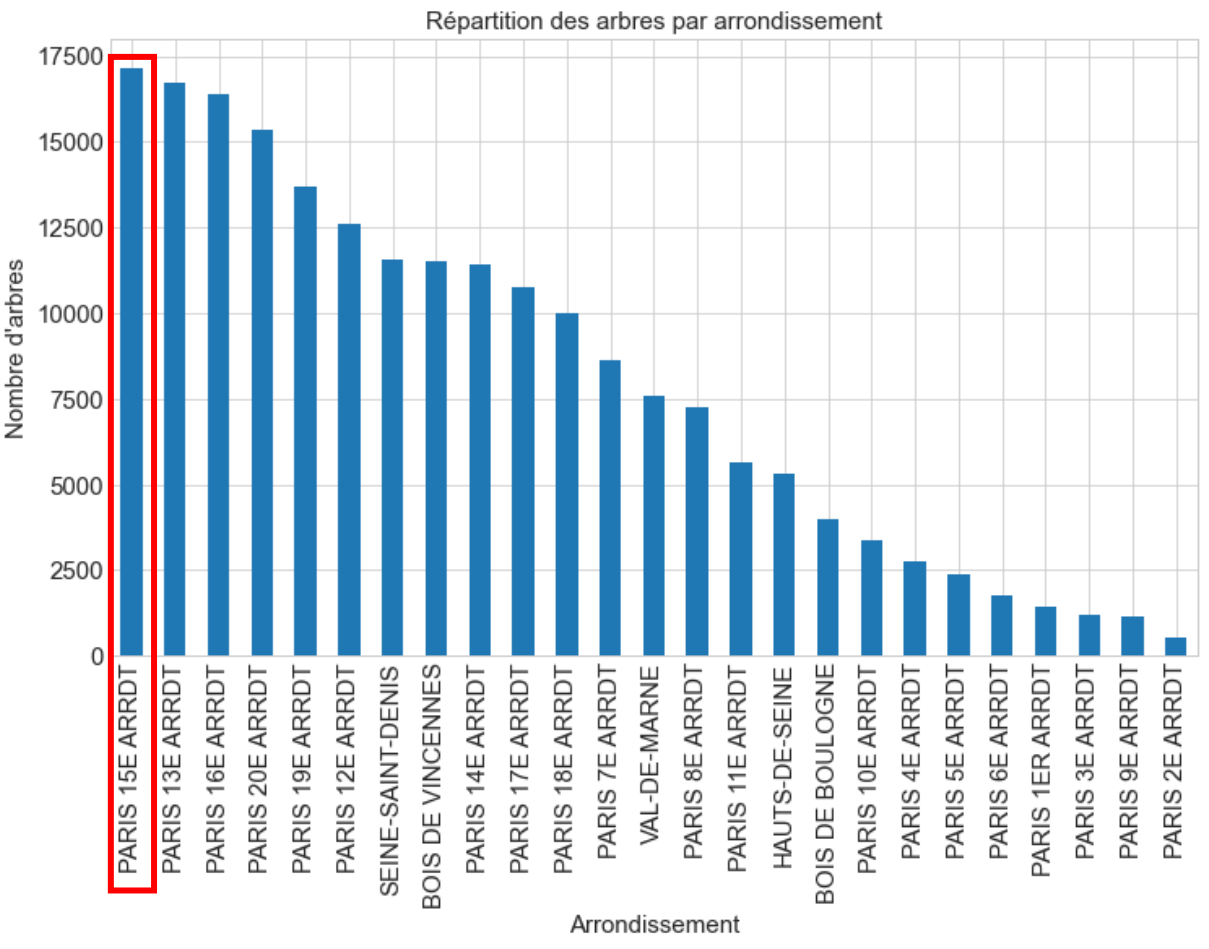
	domanialite	id_emplacement	lieu	genre	variete	libelle_francais	geo_point_2d_a	geo_point_2d_b	stade_developpement
13	Jardin	210034	PARC FLORAL DE PARIS / ROUTE DE LA PYRAMIDE	Fagus	Rotundifolia'	Hêtre	48.836862	2.446518	JA
20	Jardin	210035	PARC FLORAL DE PARIS / ROUTE DE LA PYRAMIDE	Malus	NaN	Pommier à fleurs	48.836862	2.446518	NaN

Analyse des données statistiques



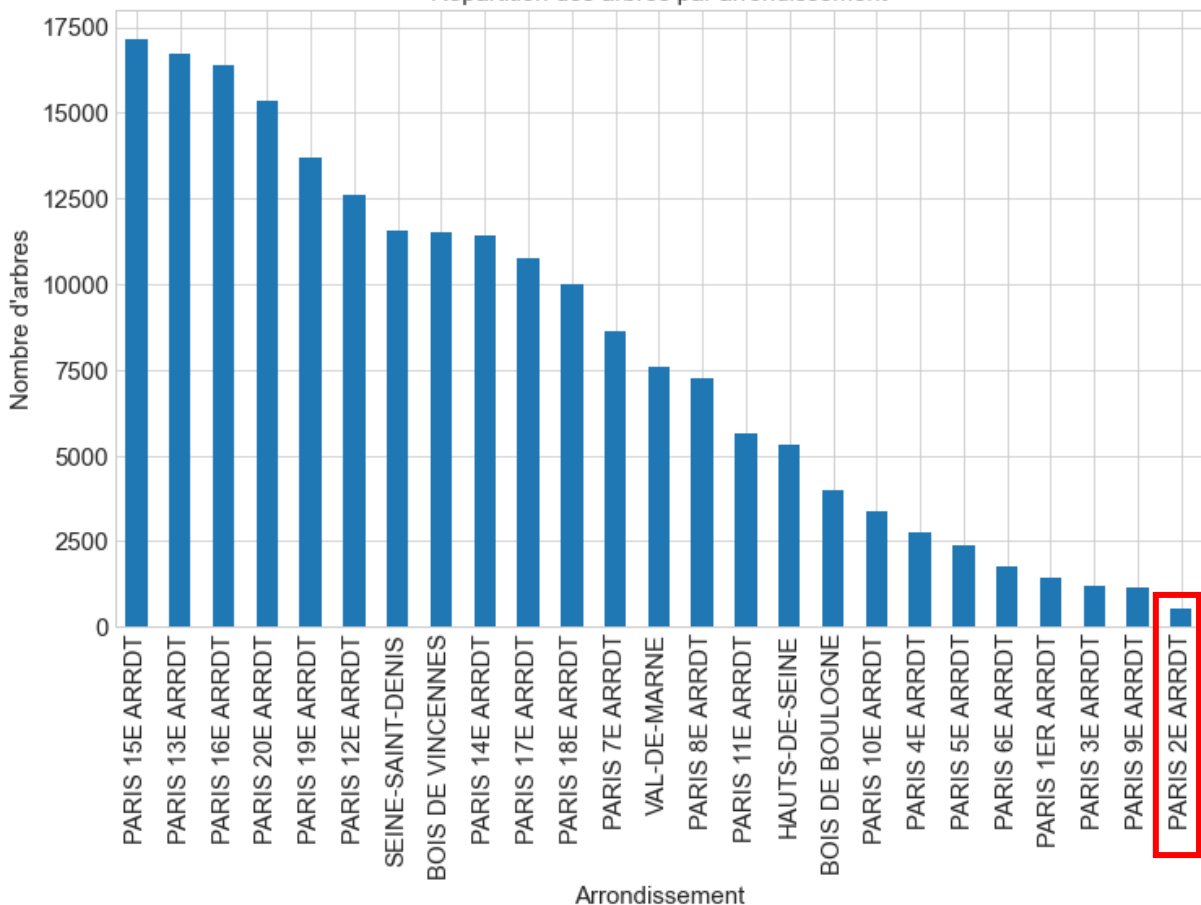
Valeurs manquantes





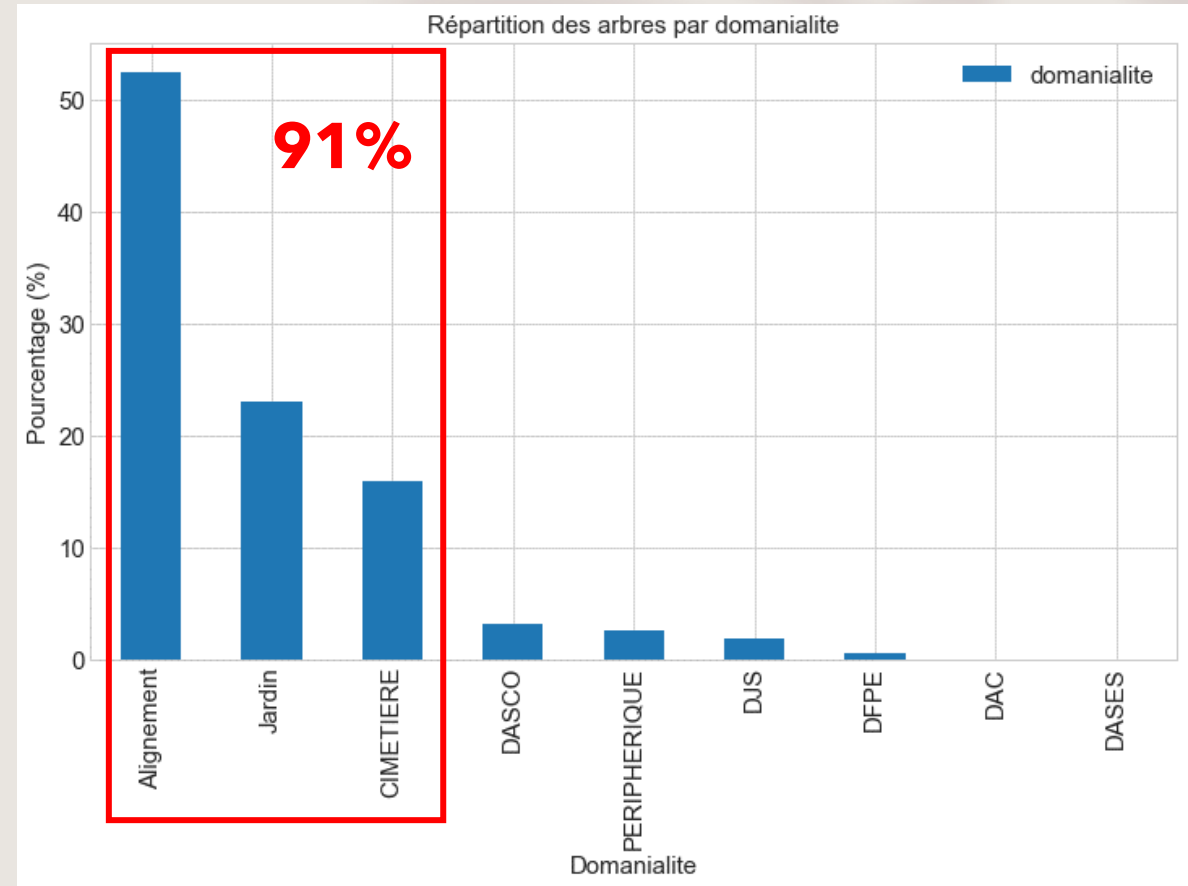
Arrondissements

Répartition des arbres par arrondissement



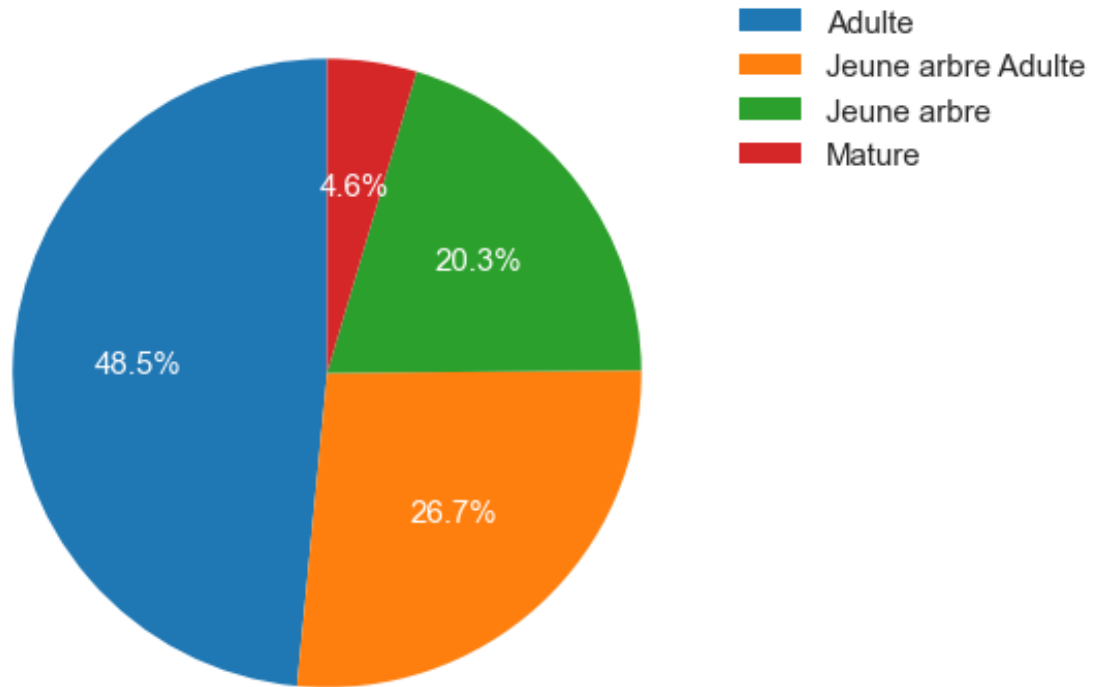
Arrondissements (suite)

Domanialite

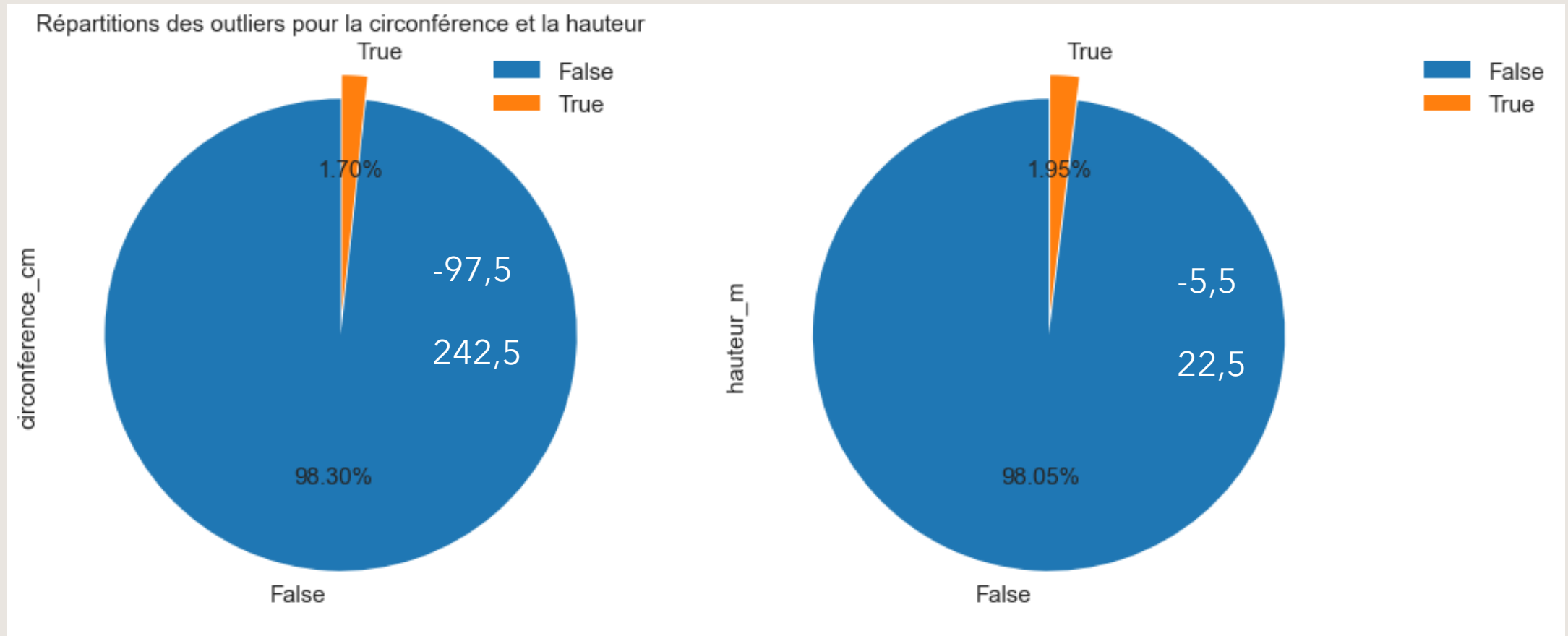


Stade de développement

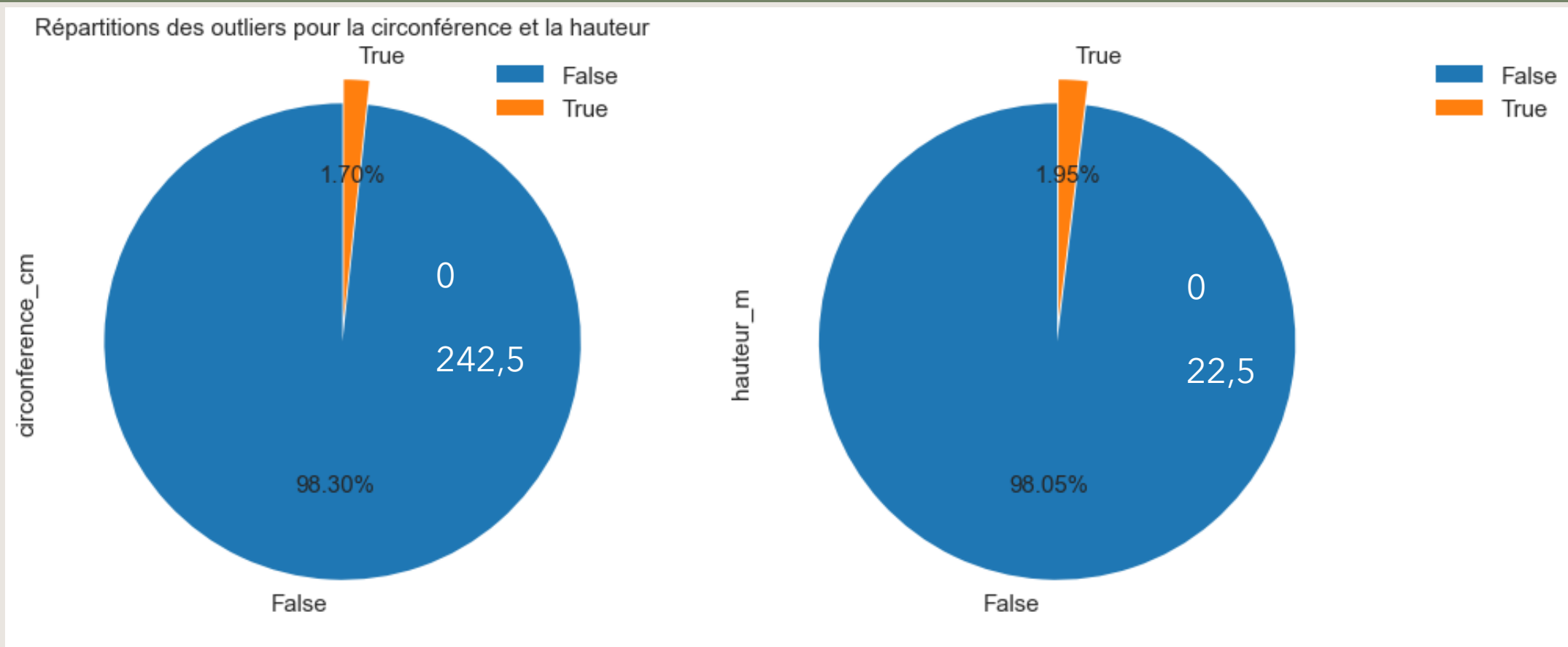
Répartition des stades de développement



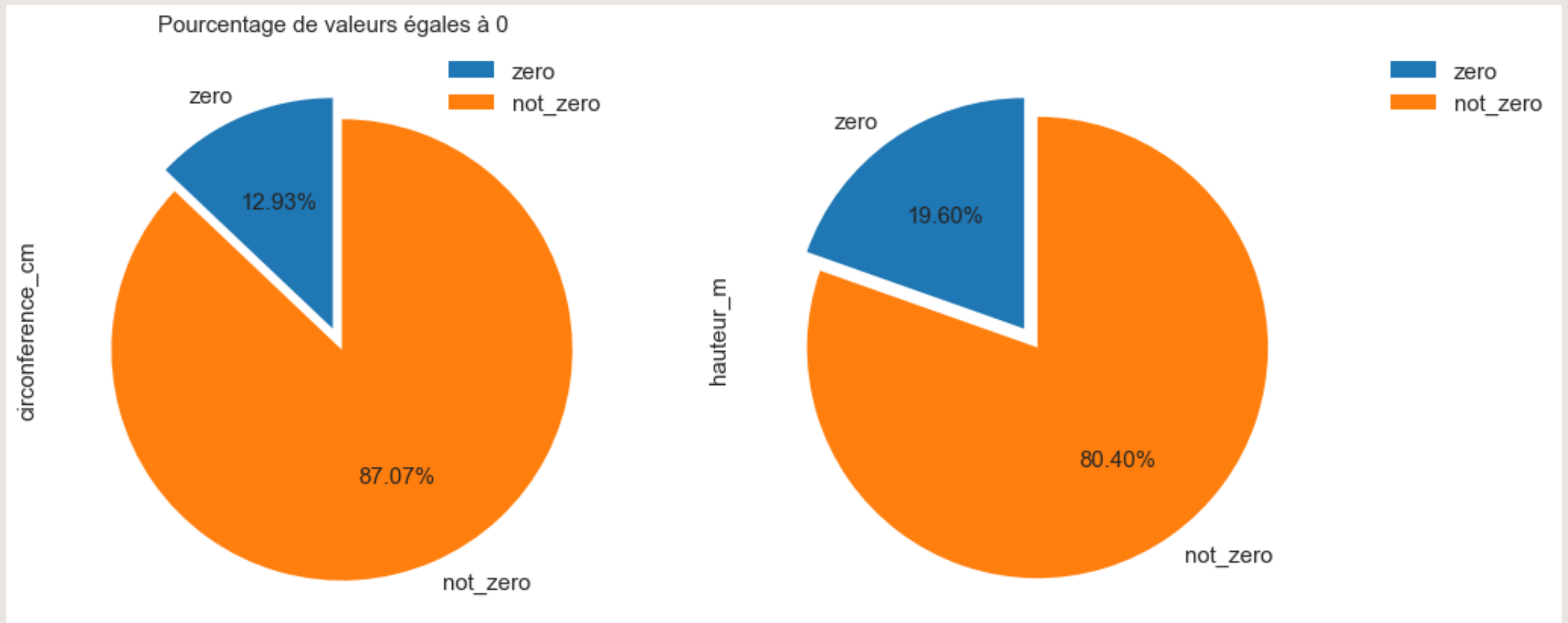
Analyse des outliers (multiple de quantiles)



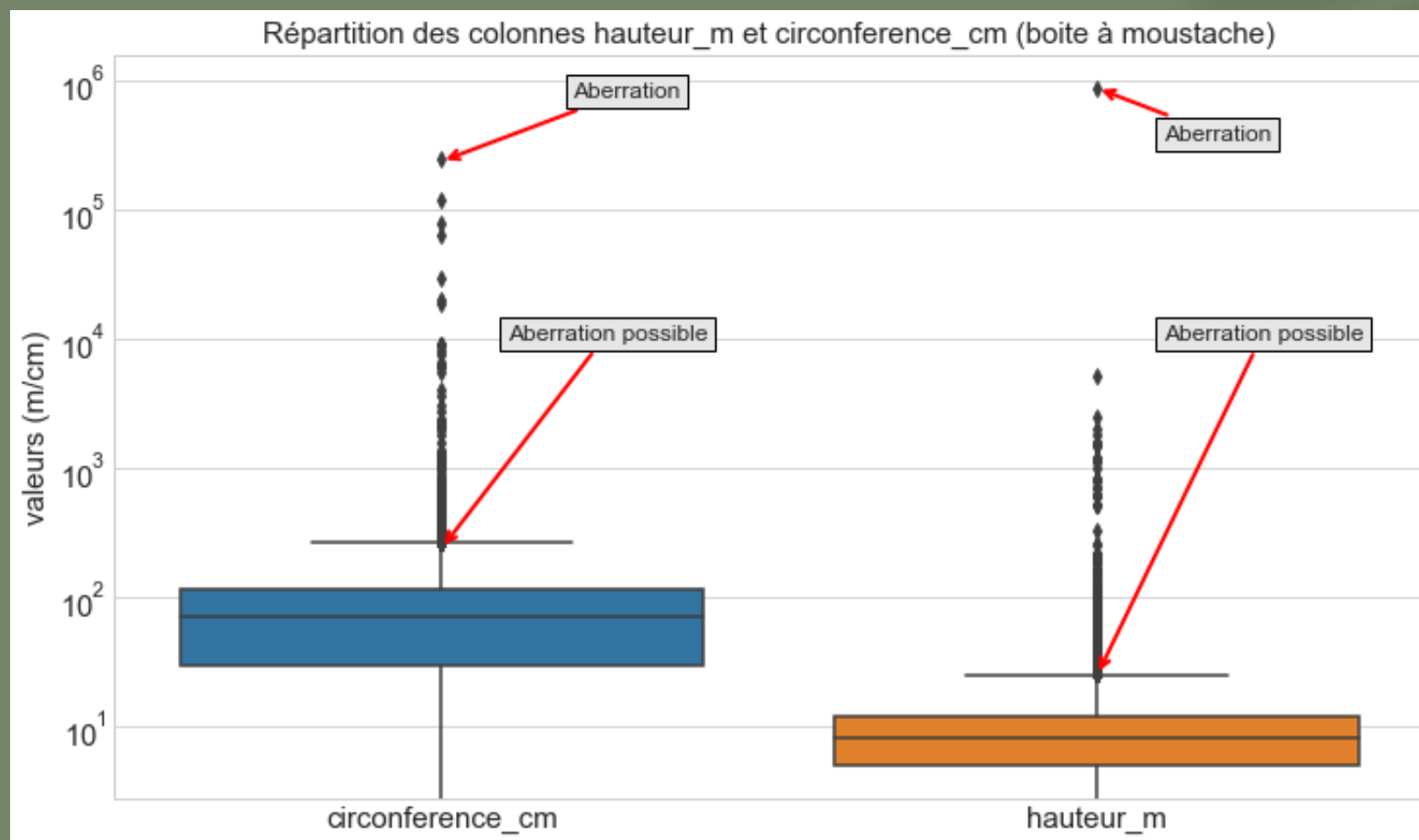
Analyse des outliers (multiple de quantiles) (suite)



Analyse valeurs égales à 0

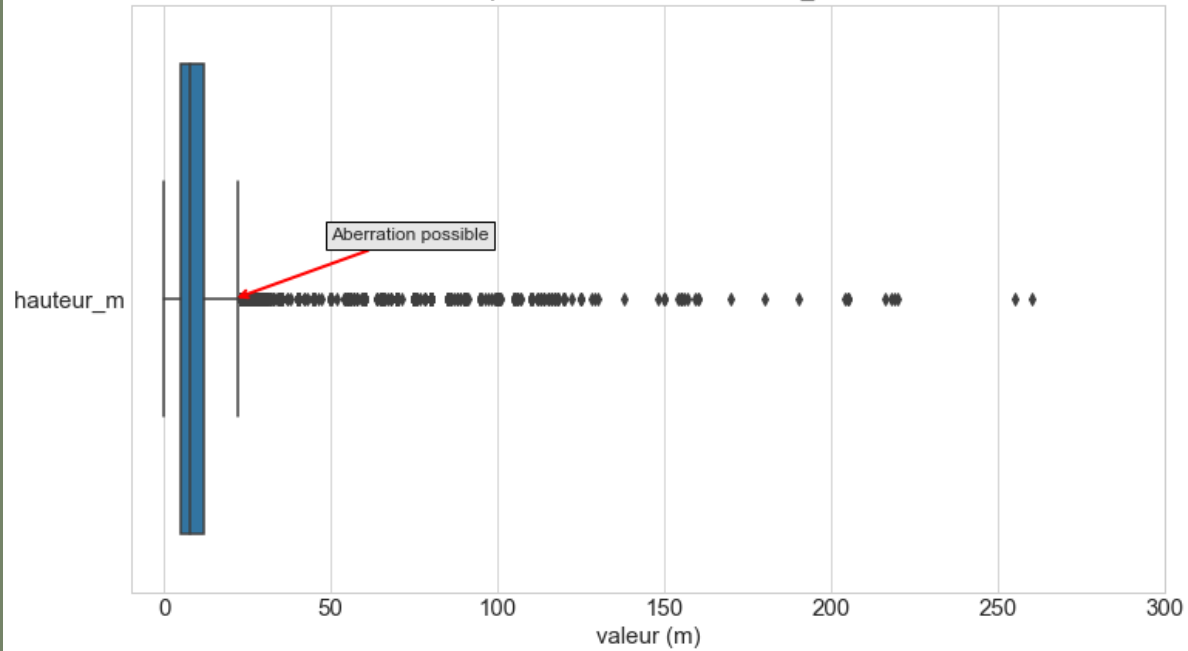


Analyse des aberrations

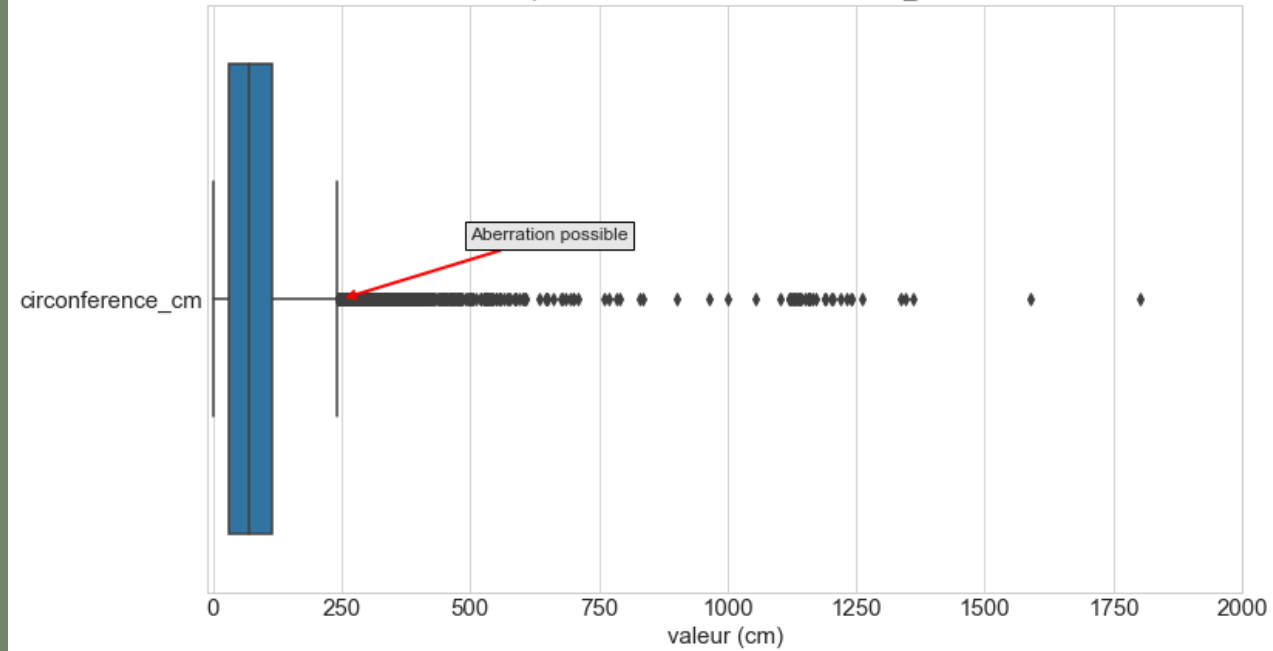


Analyse des aberrations (suite)

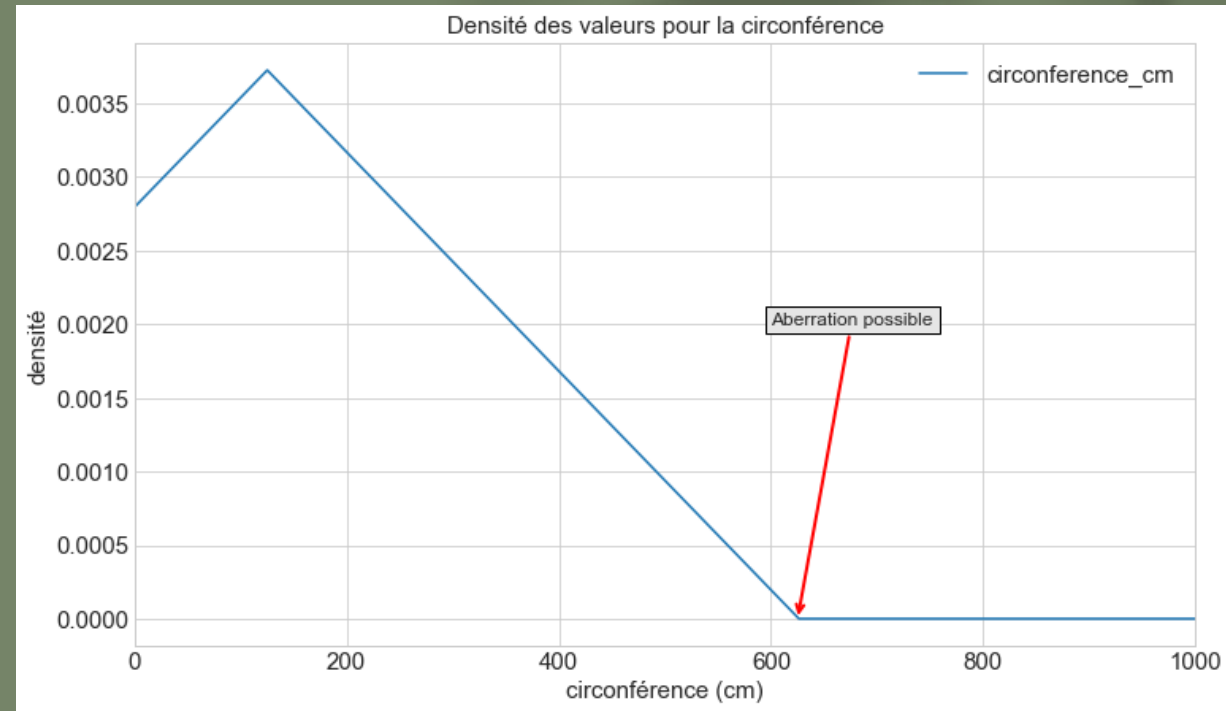
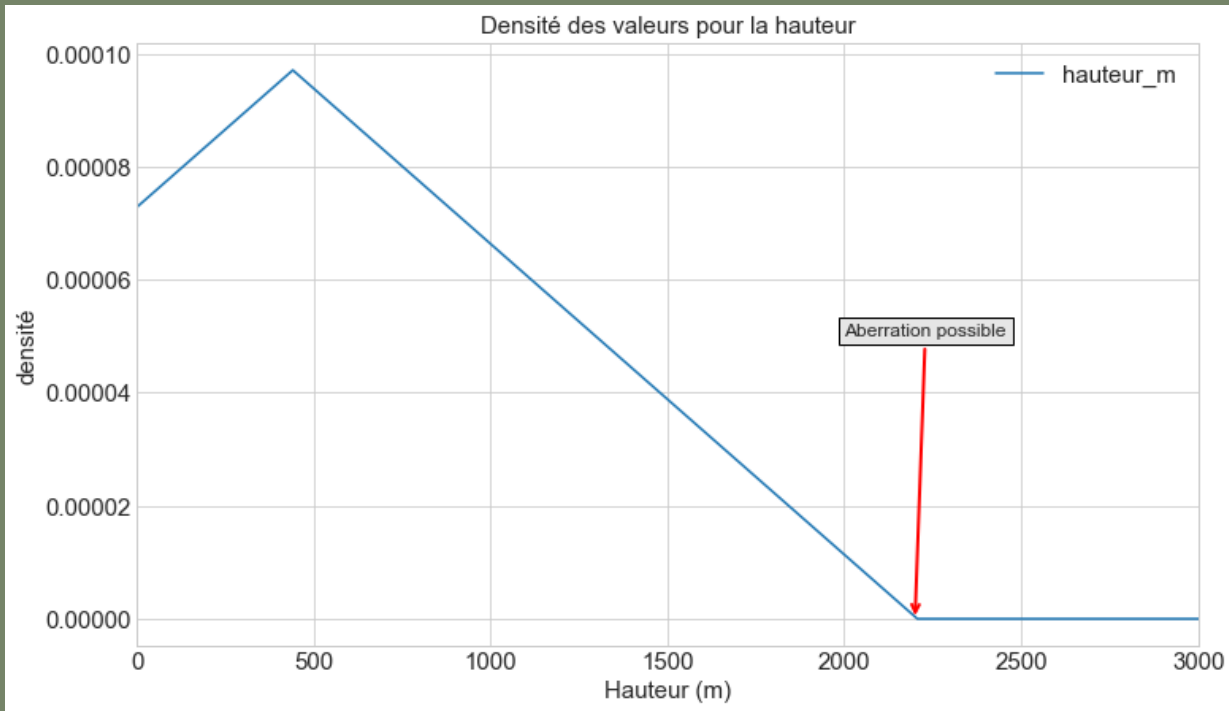
Répartition de la colonne hauteur_m



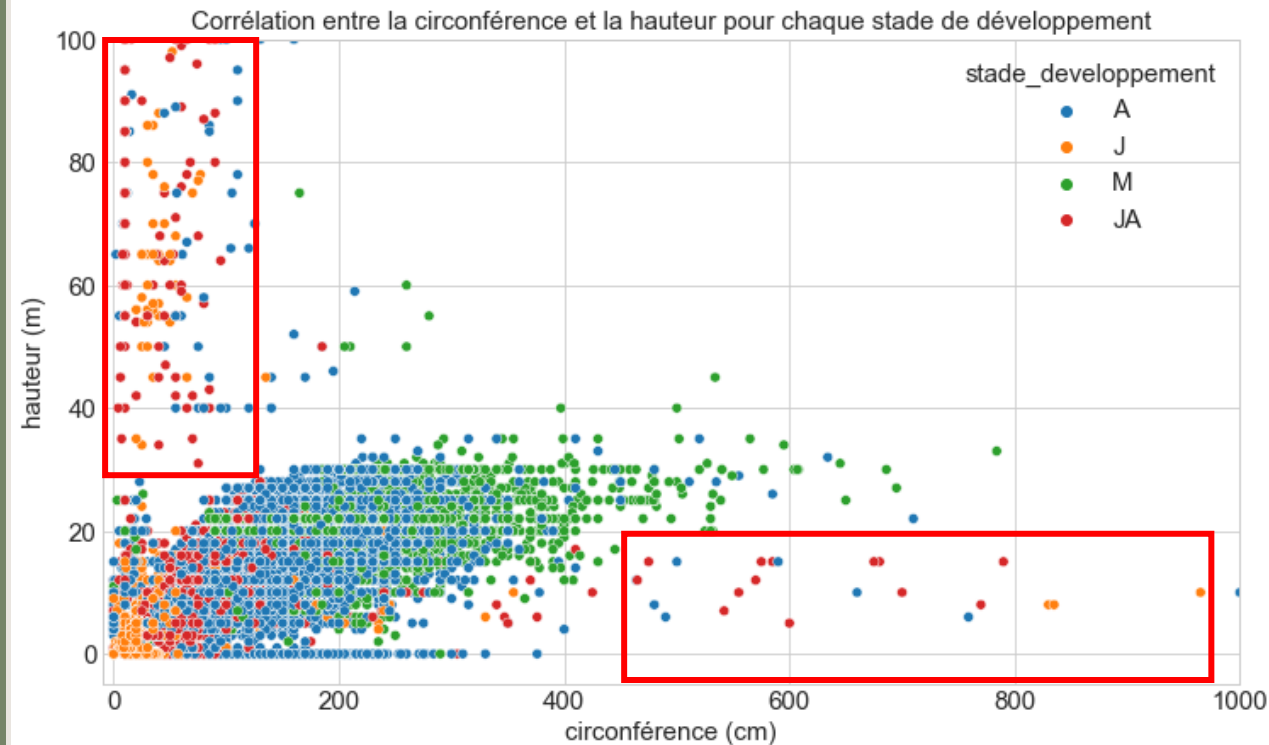
Répartition de la colonne circonference_cm

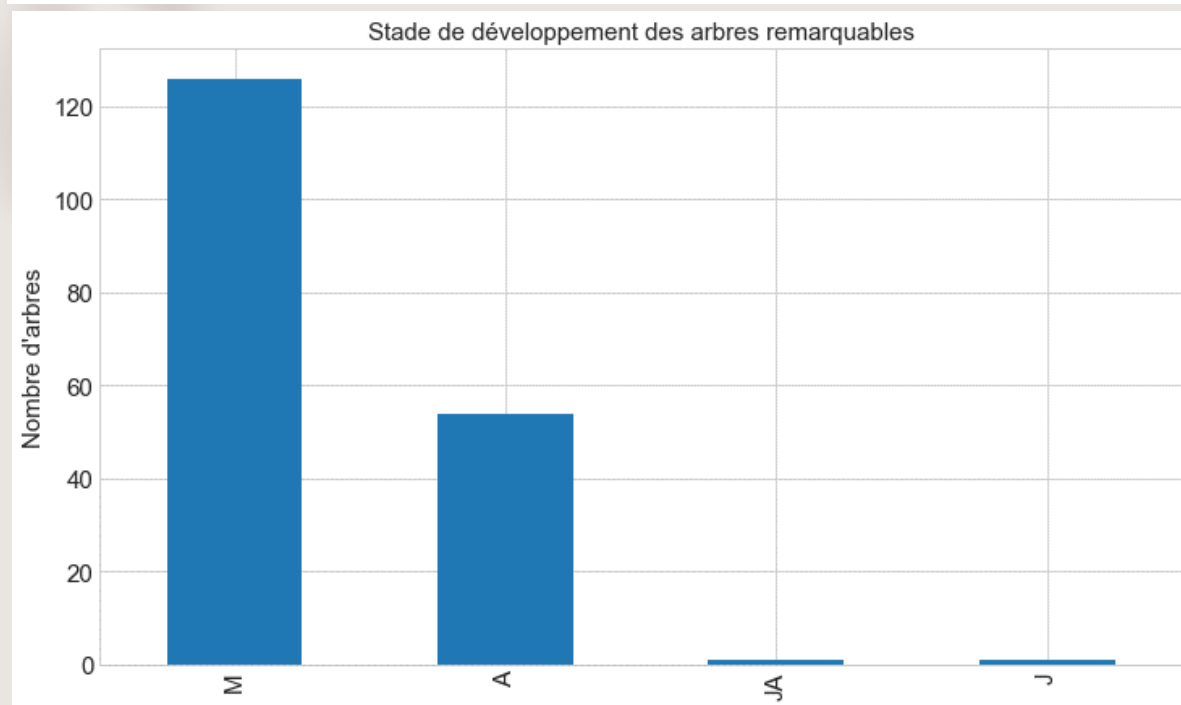
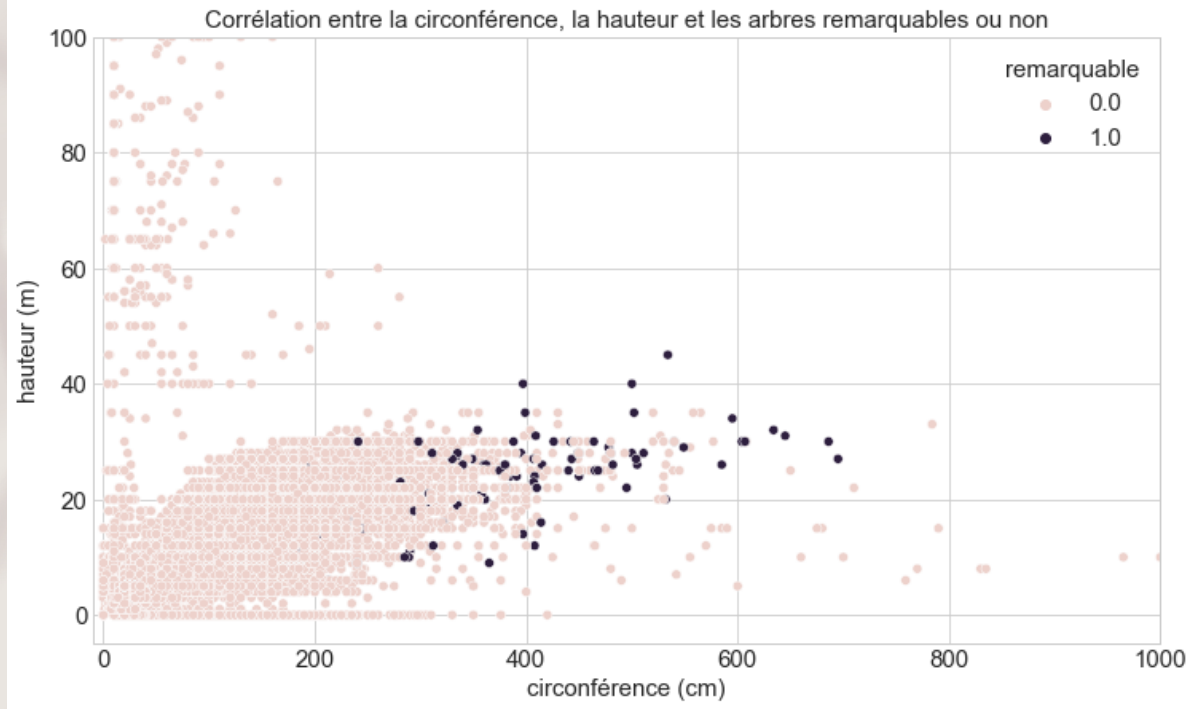


Analyse des aberrations (suite)



Circonférence / hauteur par stade développement





Remarquable /
stade de
développement

Arbres allergènes

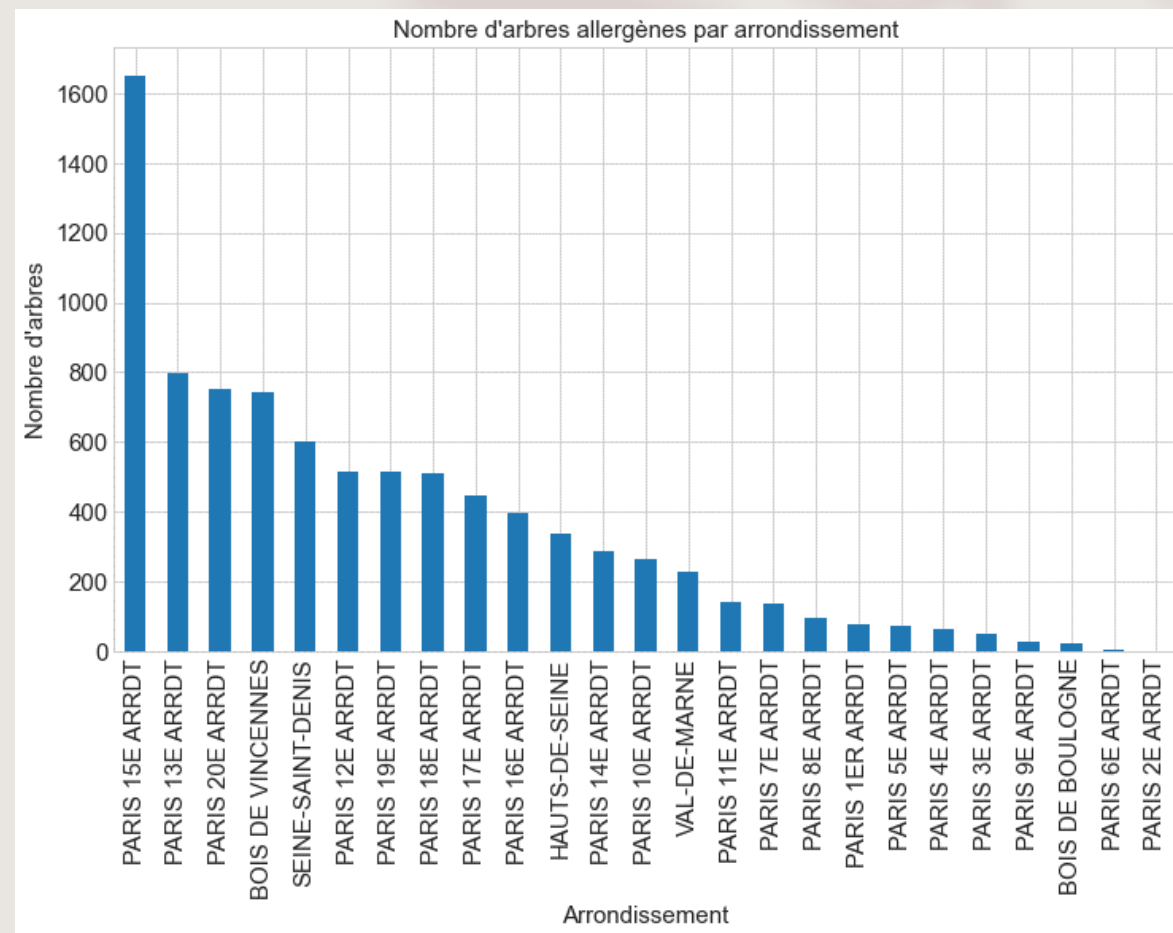
Le Saule Marsault

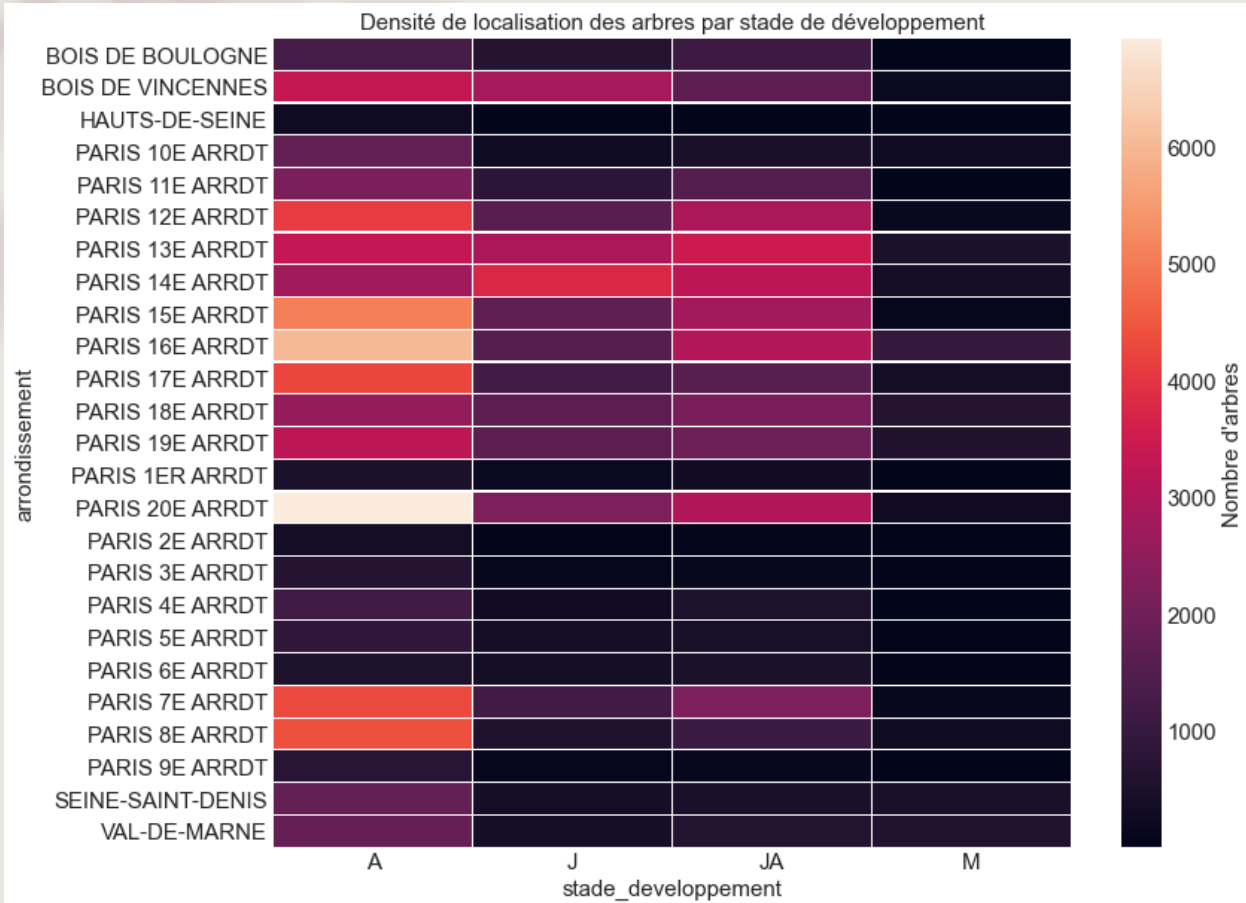
L'Aulne Glutineux

Le Bouleau Pleureur

Le Charme Commun

L'Hêtre Commun



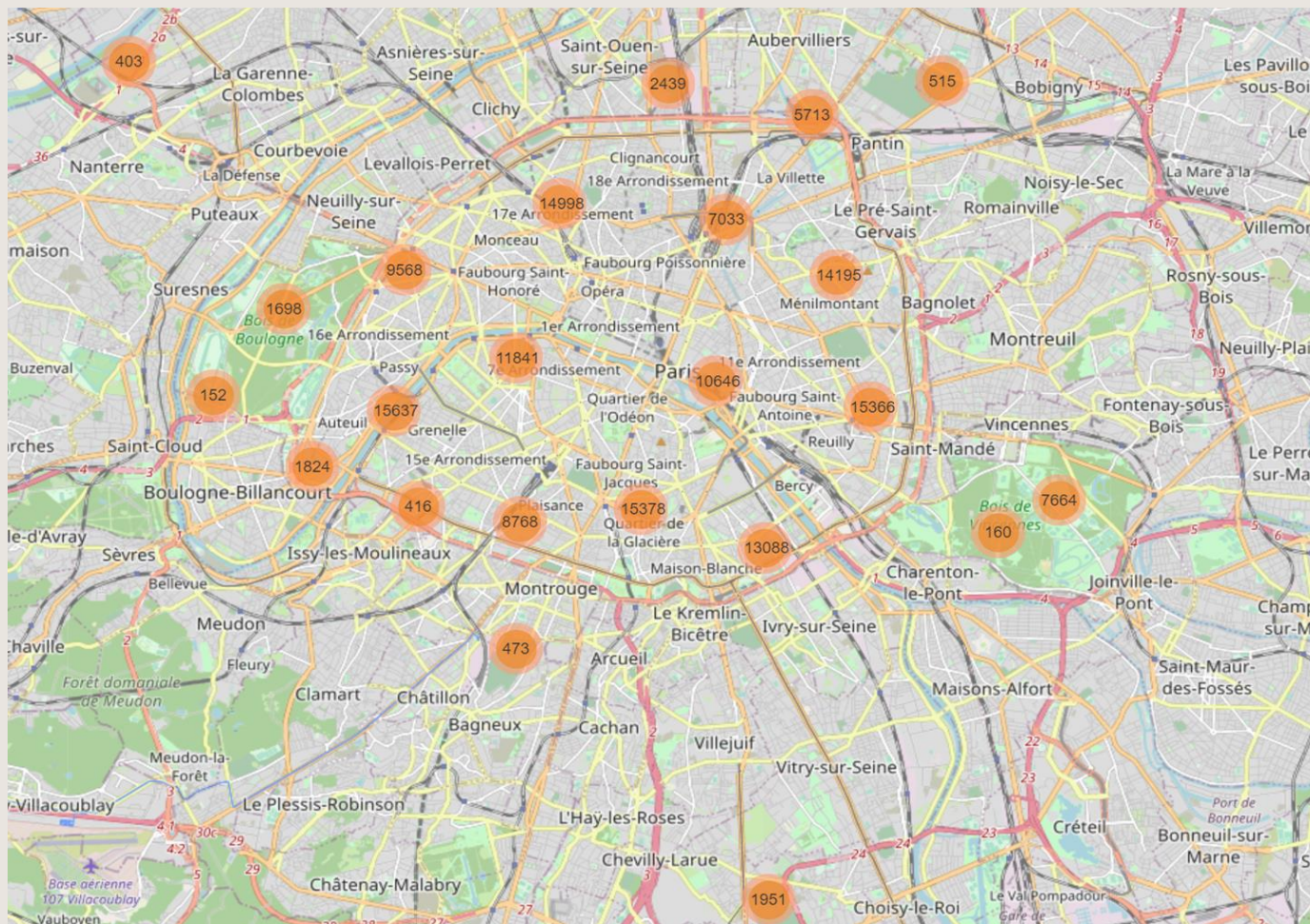


Répartition par
arrondissement /
stade de
développement

Synthèse



Synthèse



1

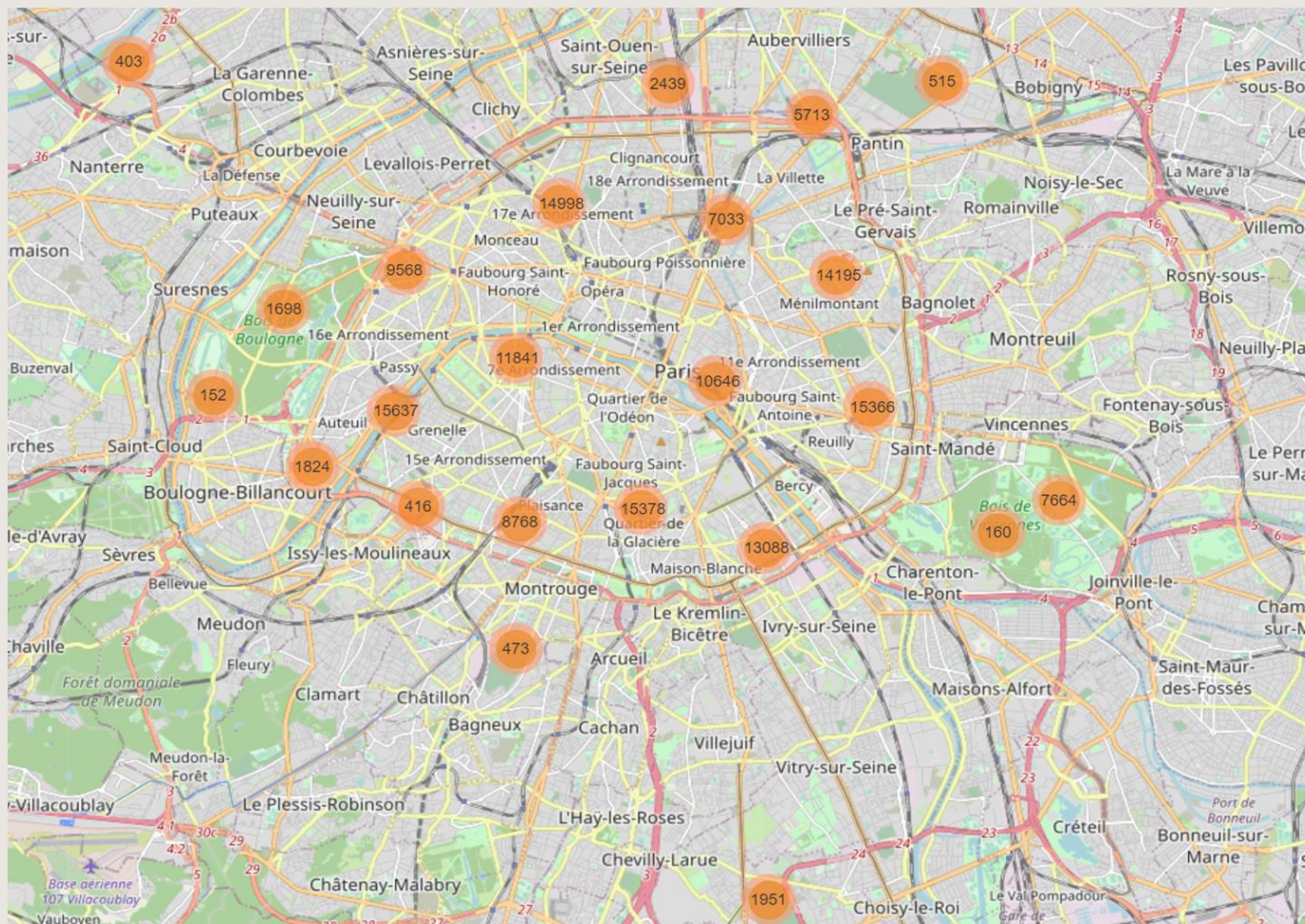
Première vraie
analyse
exploratoire

Temps passé à
comprendre la
structure



Perte de temps

Synthèse (suite)



Analyse Importante

Possible d'aller plus loin



Fournit beaucoup d'information

Intégrer l'IA pour ajouter une nouvelle dimension



A photograph of a dense forest with tall, slender trees and a thick canopy of bright green leaves. Sunlight filters through the trees, creating dappled light on the forest floor. A semi-transparent grey rectangle is overlaid on the left side of the image, containing the text "Merci de votre attention !".

Merci de votre
attention !