

PROJET 2

PARTICIPEZ À UN CONCOURS SUR LA SMART CITY:

REALISER UNE ANALYSE EXPLORATOIRE EN PYTHON D'UN JEU DE DONNEES

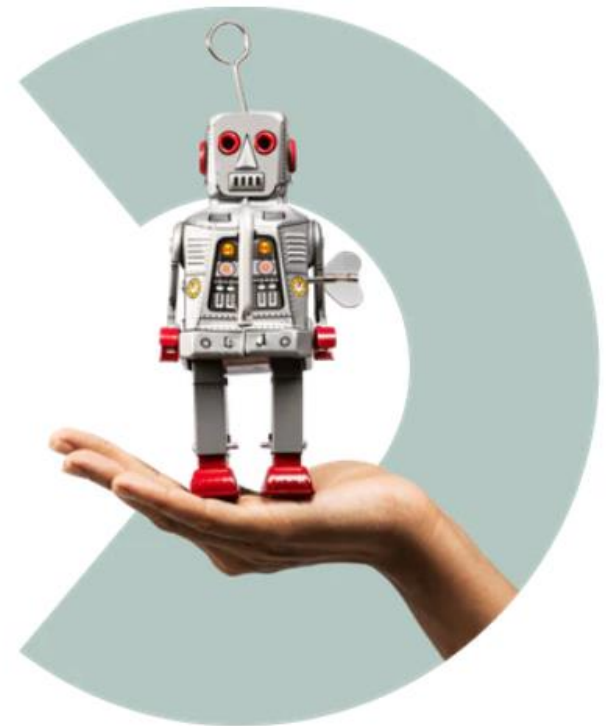
- **PORTANT SUR LES ARBRES DE PARIS**
- **DE FACON À OPTIMISER LES TOURNEES**

**#ANALYSE EXPLORATOIRE
#PYTHON #PANDAS #MATPLOTLIB
#CLEANING #FILTRE**

Ingénieur IA

Développez et intégrez des algorithmes de Deep Learning au sein d'un produit IA

Ouddane Nabil



SOMMAIRE

Projet 2

Participez à un concours de la smart city

A. INTRODUCTION

1. Contexte
2. Objectifs

B. PREREQUIS AU PROJET

1. Python
2. Données: Les arbres de Paris
3. Faits: Les arbres de Paris

C. **Projet: analyse exploratoire de la base des arbres de paris**

1. Analyse du fichier de données
2. Nettoyage des données quantitatives
 - a) 1er filtre de bon sens
 - b) Filtre IQR (interquartile range) par espèces
3. Estimation des stades de development manquants
4. Analyse géographique
 - a) par arrondissement
 - b) par arrondissement et domanialité
 - c) par genre d'arbres
 - d) par stade de developpement
5. Focus sur les jeunes arbres

D. CONCLUSION

1. Les Conclusions



INTRODUCTION

1. Contexte

- **ENJEU global:** Aider Paris à devenir une smart city en participant à un challenge data dans le cadre du programme « végétalisons la ville »
- **ENJEU DU P2: Réaliser une analyse exploratoire d'un jeu de données portant sur les arbres de Paris de façon à optimiser les tournées pour l'entretien des arbres de la ville.**
- **DONNEES SOURCES**
 - le jeu de données des arbres de la ville de Paris:
 - <https://s3-eu-west-1.amazonaws.com/static.oc-static.com/prod/courses/files/AI+Engineer/Project+2+Participez+%C3%A0+un+concours+sur+la+Smart+City/p2-arbres-fr.csv>



2. Objectifs

- **SCRIPT**: réaliser une analyse de donnée statistique univariée en python
 - Complète
 - Pertinente avec des indicateurs statistiques simples
 - Présentable avec l'utilisation des cellules markdown des jupyter notebooks





PREREQUIS AU PROJET

1. Python

- INSTALLATION DE LA SUITE ANACONDA
 - <https://www.anaconda.com/distribution/>
- CHOIX DE L'ENVIRONNEMENT JUPYTER NOTEBOOK
 - Adapté pour les data scientist

2. Données: Les 200000 arbres de Paris

- Depuis 2014, une base de données informatisée permet de suivre en temps réel chaque arbre
 - **La base de donnée « les arbres » localise géographiquement l'ensemble des arbres, ainsi que les arbres d'alignement, présents sur le territoire parisien et des cimetières extra-muros (*hors de Paris*).**
 - **Licence:** <http://opendatacommons.org/licenses/odbl/>
 - **Producteur:** Direction des Espaces Verts et de l'Environnement - Ville de Paris
 - **Fréquence de mise à jour:** Hebdomadaire
 - **Champs notables de la base de donnée**
 - Id
 - Type d'emplacement = arbre
 - Domanialité : Alignement/ Jardin /Cimetière / Dasco (direction des affaires scolaires : ecoles, collège etc..) / Périphérique / DJS (direction de la jeunesse et des sports) / DFPE (direction des familles et de la petite enfance)/ DAC (direction des affaires culturelles) / DASES (direction de l'action sociale, de l'enfance et de la santé)
 - Arrondissement : arrondissement, bois et départements limitrophes
 - Lieu/adresse
 - Id emplacement
 - Libellé français : noms des arbres en français
 - Genre
 - Espece
 - Variété ou cultivar
 - Circonférence
 - Hauteur en m : attention aberrations
 - Stade de développement: adulte / jeune arbre / jeune arbre adulte / mature
 - Flag remarquable: oui (178) ou non
 - Coordonnées GPS
-

3. Faits: Les arbres de Paris

- Paris compte plus de **200 000 arbres** (hors bois: 300 000 autres gérés différemment) situés dans les rues, les espaces verts et les équipements municipaux
 - **Moins de 1,5 % (3000 arbres) du patrimoine arboré doit être abattu** pour des raisons de sécurité.
 - Un arbre doit être abattu lorsqu'il présente un risque de chute ou de contamination d'autres arbres.
 - De jeunes arbres issus de la pépinière de la Ville sont **systématiquement replantés**.
 - Les arbres plantés le long des rues dépassent rarement 80 ans
 - Chaque année 20 % d'entre eux font l'objet d'un relevé phytosanitaire détaillé par les agents sylvicoles.
-

3. Faits: Les arbres de Paris

- la plupart des maladies n'attaquent qu'une seule espèce.
 - C'est pourquoi la concentration d'arbres d'une même espèce sur un même lieu favorise la propagation des épidémies.
 - Les principales maladies ou parasites qui attaquent les arbres de paris sont les suivantes:
 - **Le tigre du platane**
 - **La mineuse du marronnier blanc (le rouge étant moins sensible)**
 - **La graphiose de l'orme** (les cultivars « Lutèce » et « sapporo gold » sont de nouvelles variétés résistantes à la maladie)
 - **La maladie de la suie de l'érable**
 - **Entretien:**
 - **Taille:**
 - **Taille de formation sur les jeunes arbres**
 - **Taille raisonnée, élagage tous les 7 à 9 ans des arbres adultes sur les voies publiques, taille de réduction**
 - **Taille architecturée (champs de mars, place des vosges ...)**
 - **Plantation:**
 - **3 000 arbres issus des pépinières municipales plantés chaque année**
 - Les arbres sont cultivés en pépinière en moyenne 8 à 10 ans avant d'être plantés en ville.
 - Une fois planté, des soins sont apportés à l'arbre **pendant 3 ans** pour garantir une bonne reprise et un développement harmonieux
-



PROJET: ANALYSE EXPLORATOIRE DE LA BASE DES ARBRES DE PARIS

1. Analyse du fichier de données

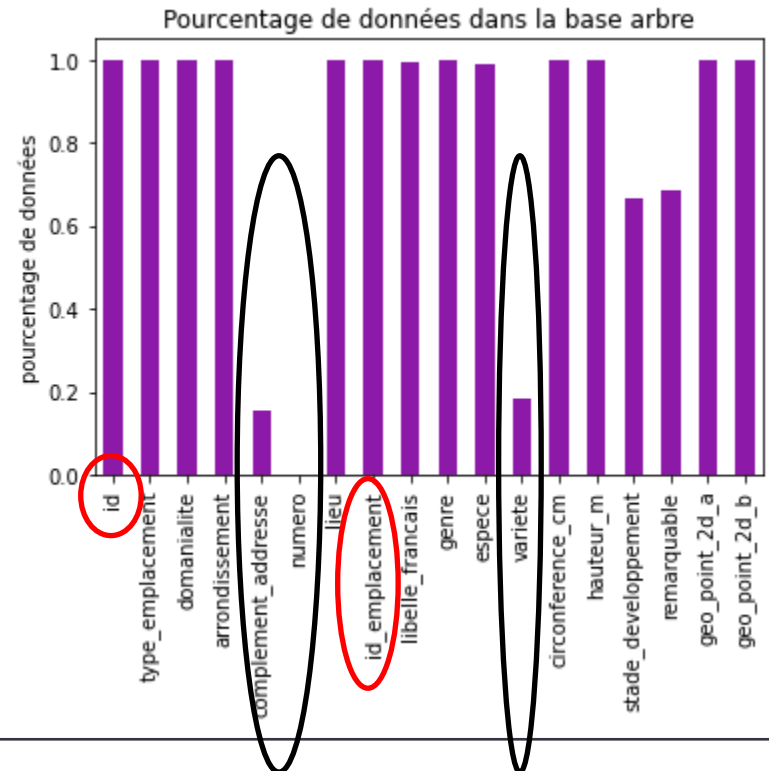
- Chargement du fichier dans un dataframe pandas
- 1ere analyse rapide:
 - **18 champs/colonnes – 200137 lignes**
 - *Les champs: complément_adresse/numero/variété peu renseignés peuvent être mis de côté faute de données*
 - *Les champs id/id_emplacement qui représentent des clés n'aideront pas notre analyse*

le fichier contient 200137 lignes

id	200137
type_emplacement	200137
domanialite	200136
arrondissement	200137
complement_adresse	30902
numero	0
lieu	200137
id_emplacement	200137
libelle_francais	198640
genre	200121
espece	198385
variete	36777
circonference_cm	200137
hauteur_m	200137
stade_developpement	132932
remarquable	137039
geo_point_2d_a	200137
geo_point_2d_b	200137

dtype: int64

test unicité des valeurs du champs id: 1.0 (1 = unique)

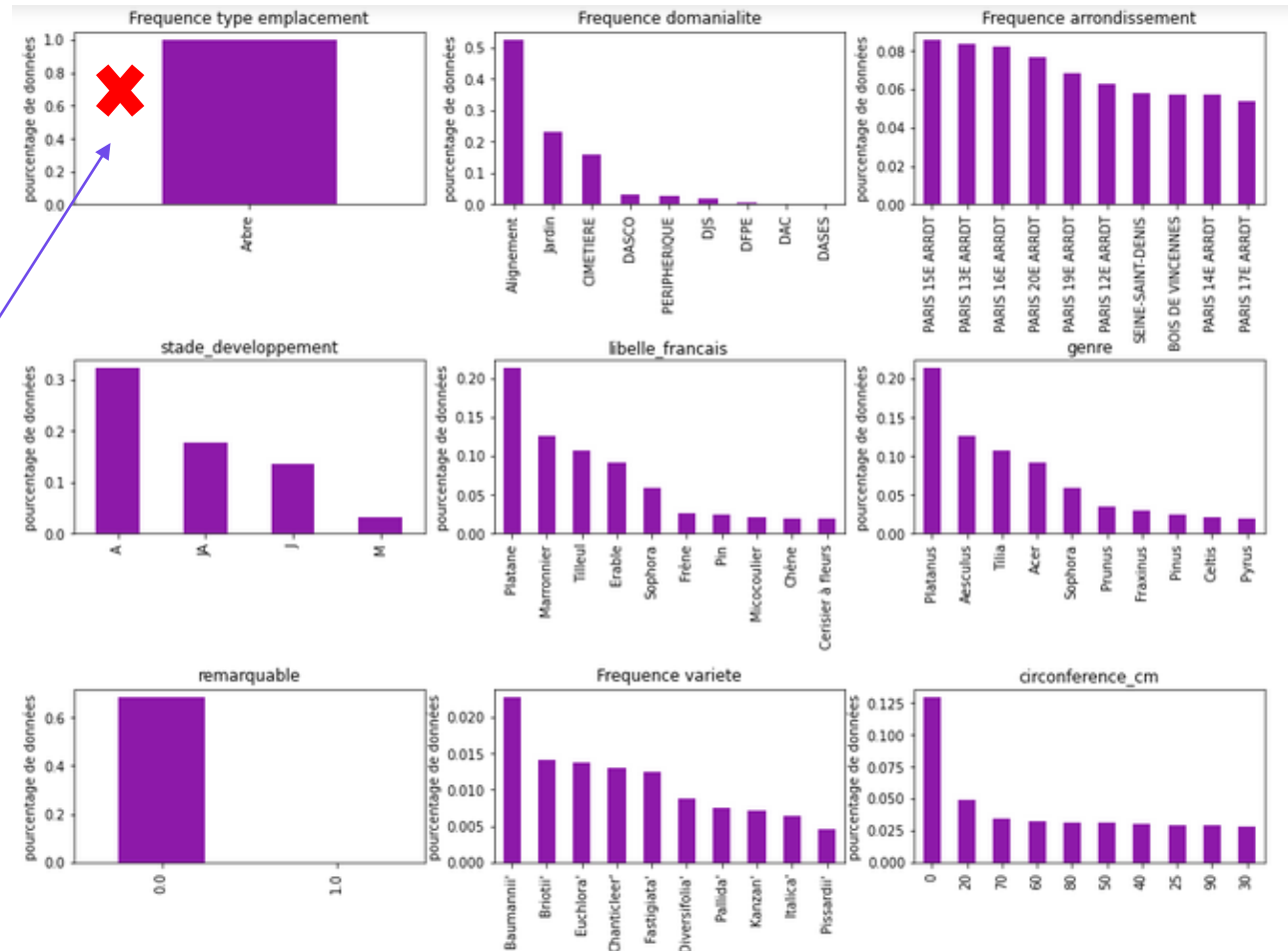


1. Analyse du fichier de données

■ Nombre de modalités par champs

```
nombre de modalites - champs type_emplacement: 1
nombre de modalites - domanialite: 9
nombre de modalites - arrondissement: 25
nombre de modalites - lieu: 6921
nombre de modalites - libelle_francais: 192
nombre de modalites - genre: 175
nombre de modalites - espece: 539
nombre de modalites - variete: 436
nombre de modalites - circonference_cm: 531
nombre de modalites - hauteur_m: 143
nombre de modalites - stade_developpement: 4
nombre de modalites - remarquable: 2
```

- Le champs type_emplacement avec une seule modalité est non discriminant donc d'aucune utilité
- Le champs lieu possède trop de modalités pour être utile dans un premier temps, le champs espèce également



1. Analyse du fichier de données

■ Résumé:

- 18 champs / 8 variables quantitatives / 10 qualitatives
- 9 champs peuvent être mis de côté pour notre analyse:
 - Les 2 champs id : id et id_emplacement
 - Le champ type_emplacement car 1 seule modalité
 - Les 3 champs de localisation par adresse: complement_adresse, numero et lieu
 - Le champs variétés qui possèdent trop de modalités et qui possèdent peu d'observations
 - Le champs remarquable: finalement trop peu d'arbres sont remarquables (moins de 0.09%)
 - Enfin le champs libellé français , qui est fortement lié au champs genre avec légèrement moins d'observations peut être abandonné. Cette liaison peut être mise en évidence de façon visuelle u par un test de χ^2 sur le tableau de contingence (liaison si χ^2 élevé et pvalue faible / attention néanmoins au nombre de degré de liberté)

χ^2 : 31976567.156921644 dl: 32661 pvalue: 0.0

	quantitative	qualitative	représentation	nb modalités	utilité	commentaires
id	x		100%	trop car unique		clé
type_emplacement		x	100%	1		1 seule et unique modalité pas discriminante
domanialité		x	~100%	ok	x	
arrondissement		x	100%	ok	x	
complement_adresse		x	~15%	trop		peu renseigné et trop de modalités
numero	x		0%			pas renseigné
lieu		x	100%	trop		trop de modalités
id_emplacement	x		100%	trop		clé
libelle francais		x	99.25%	~200	x	assez proche du genre
genre		x	~100%	~200	x	assez proche du libellé francais*
espece		x	99.12%	trop	x	trop de modalités
variete		x	~18%	trop		peu renseigné et trop de modalités
circonference_cm	x		faux 100%		x	taille à nettoyer / 0 à la place de NR / outliers
hauteur_m	x		faux 100%		x	taille à nettoyer / 0 à la place de NR / outliers
stade_developpement		x	66%	ok	x	possibilité d'augmenter sa representation
remarquable	x		68%	2		booleen / tres peu d'arbres remarquables
geo_point_2d_a	x		100%		x	coordonnées
geo_point_2d_b	x		100%		x	coordonnées

2. Nettoyage des données quantitatives

a) 1er filtre de bon sens

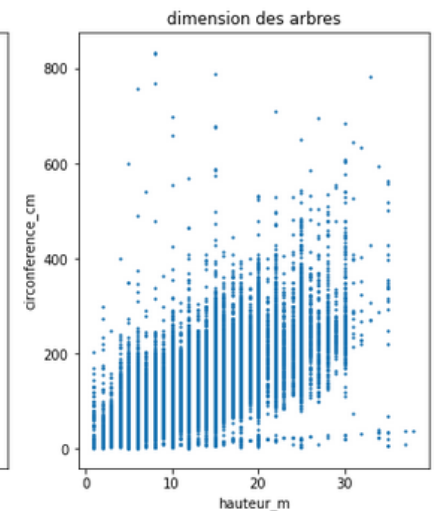
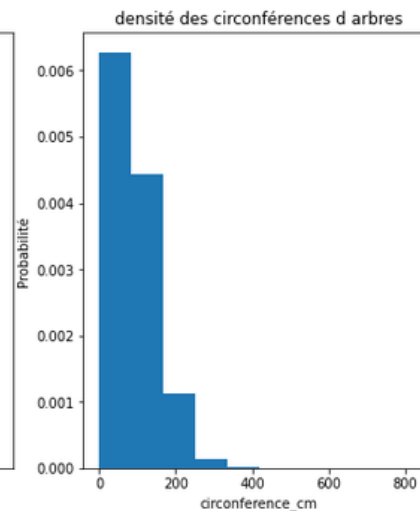
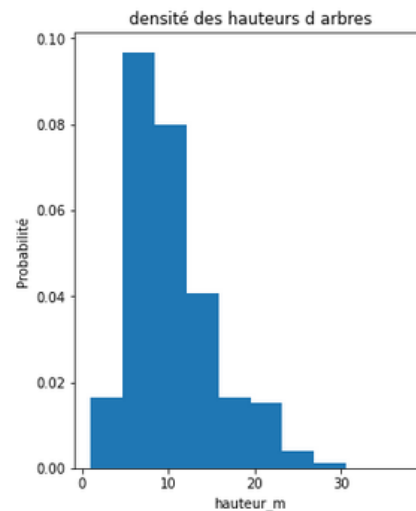
- 1^{er} filtre sur les champs de dimension des arbres
 - D'après la ville de paris , son arbre le plus haut ne dépasse pas les 35/40 m et son arbre le plus gros ne dépasse pas les 850/900 cm de circonférence au tronc. Ces valeurs peuvent constituer un premier filtre sur les valeurs aberrantes max.
 - Par ailleurs, par définition , les dimensions doivent être supérieures à 0
 - Enfin toutes les valeurs égales à 0 des champs hauteur et circonférence correspondent à des valeurs non renseignées et doivent être mises de côté et remesurées

nombre de circonférence à 0: 25867 , équivalent à 12.92 % de la table
nombre de hauteur à 0: 39219 , équivalent à 19.60 % de la table
nombre de circonférence à 0 et nombre de hauteur à 0: 25501 , équivalent à 12.74 % de la table

nombre de circonférence > 9m: 75 , équivalent à 0.04 % de la table
nombre de hauteur > 40m: 506 , équivalent à 0.25 % de la table
nombre de circonférence > 9m et nombre de hauteur > 40m: 2 , équivalent à 0.00 % de la table

min circonference: 1.0 cm
max circonference: 835.0 cm
moyenne circonference: 91.58 cm
mediane circonference: 80.0 cm
ecart type circonference: 58.95 cm
quantile 99.9% circonference: 400.0 cm
quantile 0.1% circonference: 10.0 cm

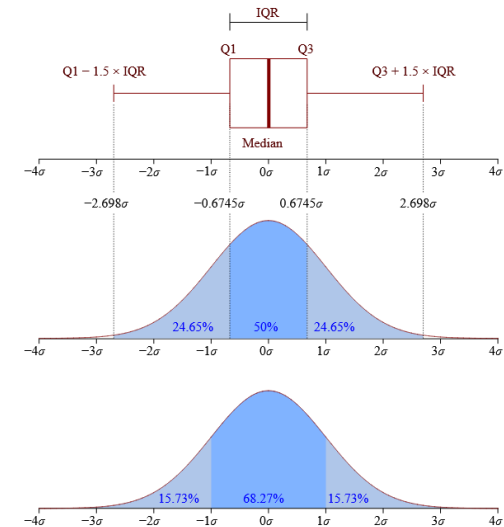
min hauteur: 1.0 m
max hauteur: 38.0 m
moyenne hauteur: 10.36 m
mediane hauteur: 10.0 m
ecart type hauteur: 5.14 m
quantile 99.9% hauteur: 30.0 m
quantile 0.1% hauteur: 1.0 m



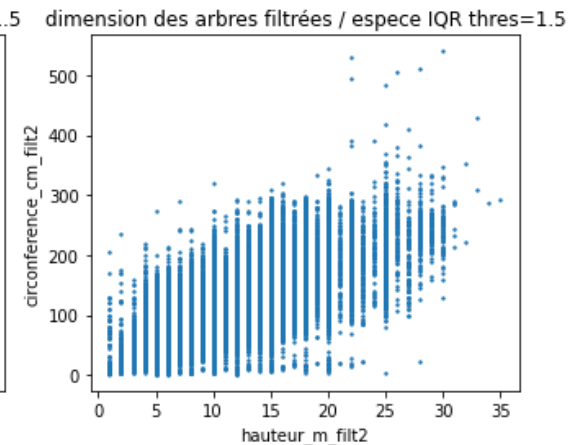
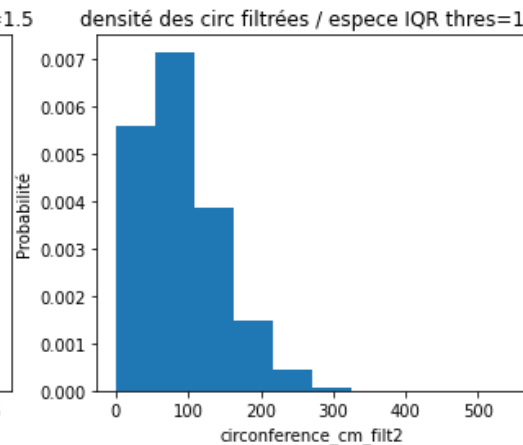
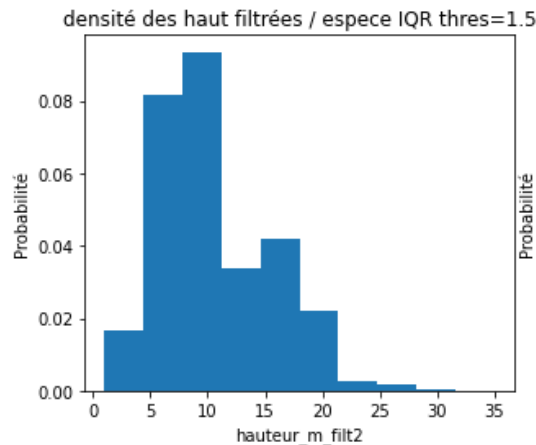
2. Nettoyage des données quantitatives

b) Filtre IQR (interquartile range) par espèces(539 mod)

- Utilisation de la méthode IQR basée sur une mesure de dispersion utilisant les quartiles.
- Par défaut le seuil/coefficient utilisé est 1.5 mais il reste paramétrable



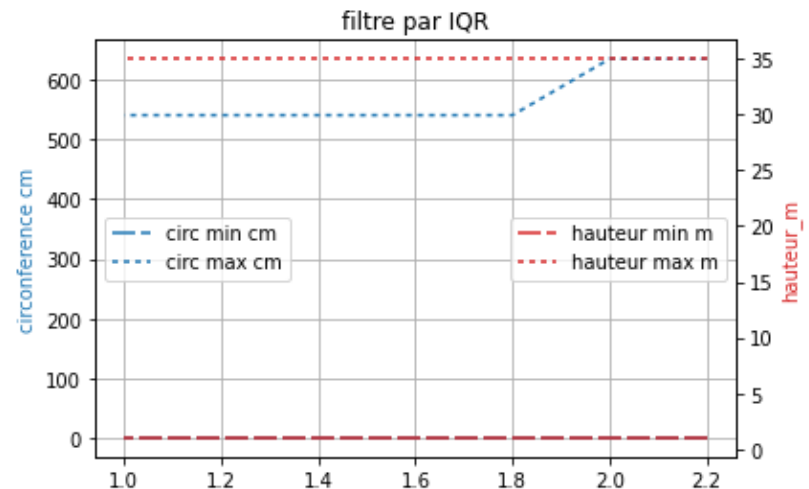
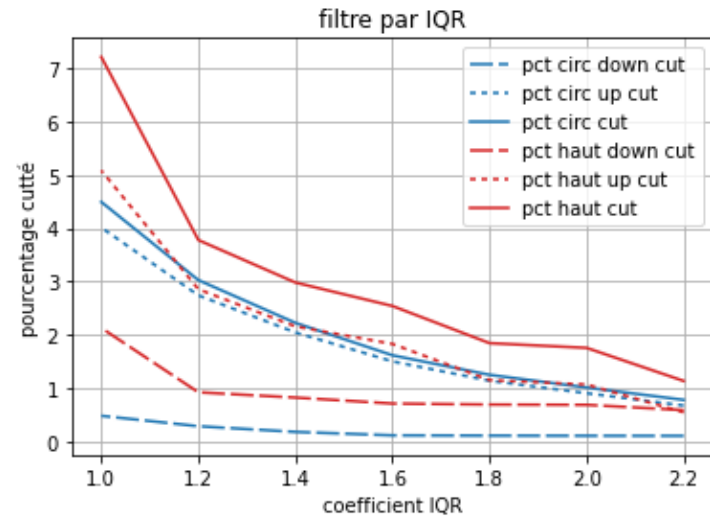
IQR threshold= 1.5 - filtration: circonférences: 1.96 % hauteurs: 2.76 % circonférences et hauteurs: 0.59 %



2. Nettoyage des données quantitatives

b) Filtre IQR (interquartile range) par espèces

- Nettoyage de données par filtre IQR
 - Coefficient seuil par défaut de 1.5
 - Test de plusieurs coefficients :
 - À coefficient égal , autant de circonférences élevées que de hauteurs hautes sont filtrées
 - A coefficient égal , peu de circonférences et hauteurs faibles sont filtrées. Ce pourcentage reste relativement stable au dessus de 1.2:
 - Les faibles valeurs ne sont pas forcément des données aberrantes, arbre jeune par exemple



3. Estimation des stades de developpement manquants

- 33.9% des stades de développement ne sont pas renseignés
- Après le filtre IQR: 13.32% des arbres ne peuvent recevoir une estimation de leur stade de développement faute de dimension renseignée
- 6% des arbres ont un stade estimable seulement par la circonférence
- 15% des arbres ont un stade de développement estimable par la circonférence et la hauteur
- Tilia, Platanus, Acer, Aesculus, Prunus représentent 46% des genres d'arbres à stade manquant => analysons plus précisément ces genres

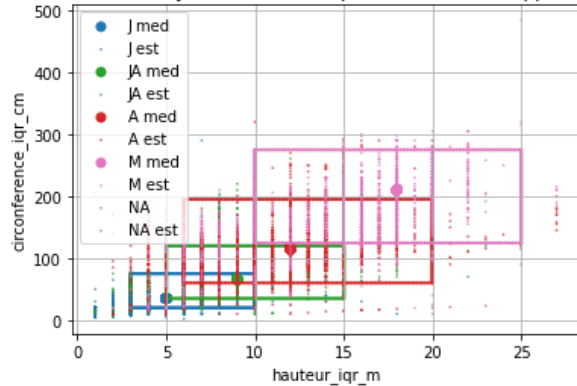
```
Pourcentage d arbres aux stades de developpement non renseignés mais estimables
Tilia      0.111939
Platanus   0.101396
Acer       0.098194
Aesculus   0.092329
Prunus     0.048179
Name: genre, dtype: float64
```

3. Estimation des stades de developpement manquants

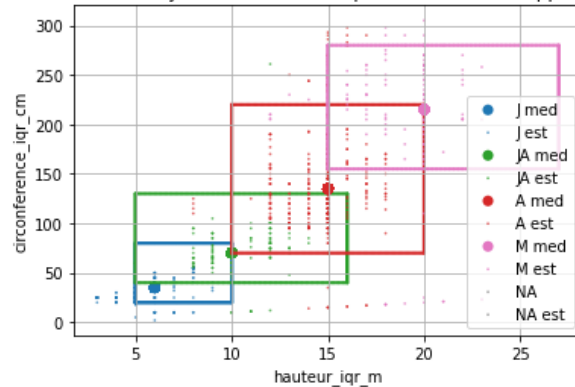
Classement supervisé

- Mise en place d'un **algorithme d'affectation de stade** basé sur la fonction **percentileofscore** (pourcentage du score dans sa distribution):
 - Pour chaque arbre, on attribue un pourcentage de score en fonction de sa hauteur , de sa circonférence ou des 2 si disponibles pour chaque stade et on attribue le stade qui a le pourcentage le plus proche de 50 (en gros le stade le plus probable)

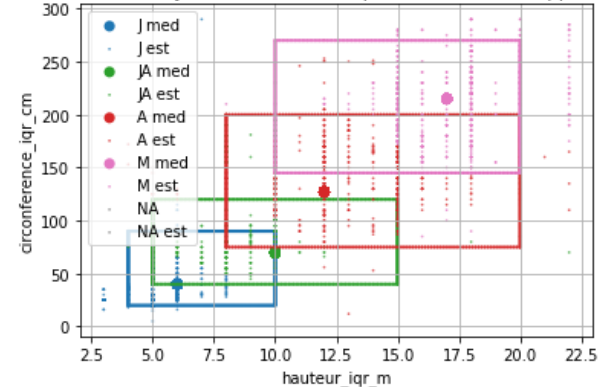
dimension moyenne des arbres par stade de developpement



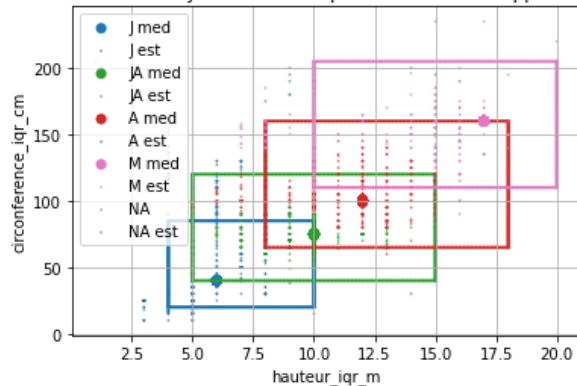
dimension moyenne des Platanus par stade de developpement



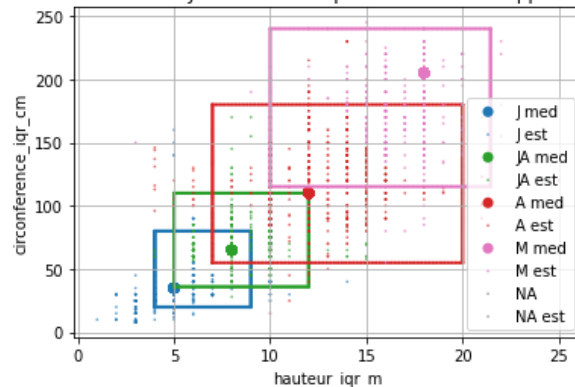
dimension moyenne des Aesculus par stade de developpement



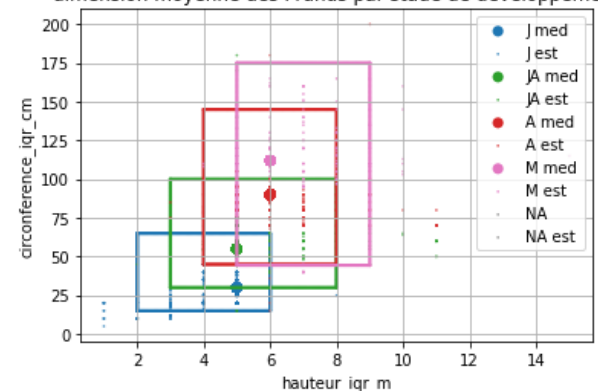
dimension moyenne des Tilia par stade de developpement



dimension moyenne des Acer par stade de developpement



dimension moyenne des Prunus par stade de developpement

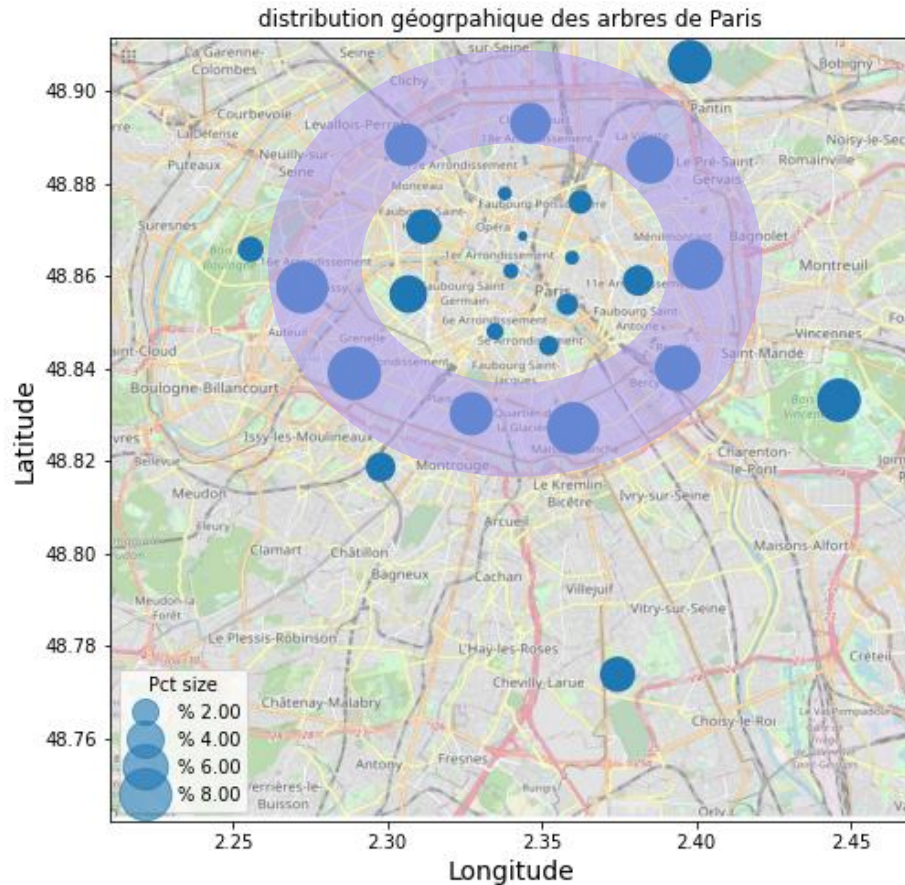


- Les carrés représentent les quantiles 5/95% des dimensions par stade (Jeune/JeuneAdulte/Adulte/Mature). Les Balises représentent les médianes

4. Analyse géographique

a) par arrondissement

- La majorité des arbres se trouvent dans les arrondissements périphériques de paris. On peut y ajouter les bois de vincennes et le 93.

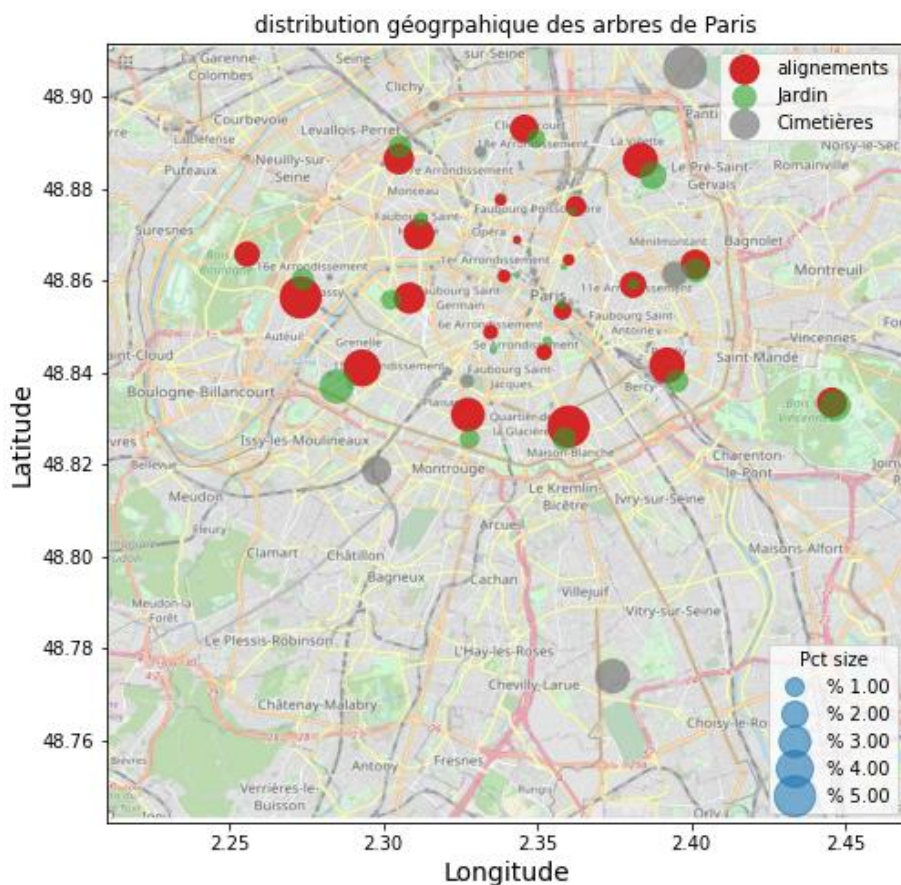


	arrondissement	pct_arbres
0	PARIS 15E ARRD	0.085566
1	PARIS 13E ARRD	0.082329
2	PARIS 16E ARRD	0.081849
3	PARIS 20E ARRD	0.076588
4	PARIS 19E ARRD	0.068363
5	PARIS 12E ARRD	0.062742
6	SEINE-SAINT-DENIS	0.057780
7	BOIS DE VINCENNES	0.057401
8	PARIS 14E ARRD	0.056876
9	PARIS 17E ARRD	0.053723
10	PARIS 18E ARRD	0.049881
11	PARIS 7E ARRD	0.042731
12	VAL-DE-MARNE	0.037839
13	PARIS 8E ARRD	0.036185
14	PARIS 11E ARRD	0.028216
15	HAUTS-DE-SEINE	0.026472
16	BOIS DE BOULOGNE	0.019861
17	PARIS 10E ARRD	0.016898
18	PARIS 4E ARRD	0.013666
19	PARIS 5E ARRD	0.011747
20	PARIS 6E ARRD	0.008769
21	PARIS 1ER ARRD	0.007025
22	PARIS 3E ARRD	0.006041
23	PARIS 9E ARRD	0.005831
24	PARIS 2E ARRD	0.002728

4. Analyse géographique

b) par arrondissement et domanialité

- Plus de la moitié des arbres sont sur des alignements, $\frac{1}{4}$ sont dans des jardins et $\frac{1}{6}$ dans les cimetières
- Les jardins sont assez corrélés en terme de taille et de position avec les alignements.
- Les cimetières possédant le plus d'arbres sont en dehors de Paris dans le 93/94/97. Le Père Lachaise fait exception



	domanialite	pct_arbres
0	Alignement	0.523052
1	Jardin	0.230872
2	CIMETIERE	0.159441
3	DASCO	0.031998
4	PERIPHERIQUE	0.026607
5	DJS	0.018407
6	DFPE	0.006600
7	DAC	0.000105
8	DASES	0.000020

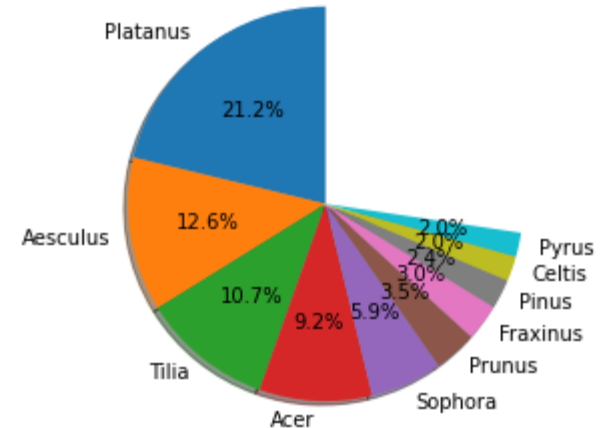
domanialite	Alignement	CIMETIERE	Jardin
arrondissement			
PARIS 13E ARRT	0.054278	0.000000	0.016833
PARIS 16E ARRT	0.053958	0.001559	0.015130
PARIS 15E ARRT	0.041686	0.000779	0.036400
PARIS 12E ARRT	0.038644	0.000035	0.016089
PARIS 19E ARRT	0.037389	0.000485	0.022789
PARIS 14E ARRT	0.035166	0.006146	0.011957
PARIS 7E ARRT	0.030149	0.000000	0.012167
PARIS 8E ARRT	0.029390	0.000000	0.006595
PARIS 17E ARRT	0.028750	0.003847	0.016324
BOIS DE VINCENNES	0.027426	0.000000	0.029974
PARIS 20E ARRT	0.026742	0.020236	0.017678
PARIS 18E ARRT	0.024433	0.004407	0.010743
PARIS 11E ARRT	0.021265	0.000000	0.004352
BOIS DE BOULOGNE	0.019851	0.000000	0.000005
PARIS 10E ARRT	0.012926	0.000000	0.002203
PARIS 4E ARRT	0.009743	0.000000	0.003243
PARIS 5E ARRT	0.007860	0.000000	0.002873
PARIS 6E ARRT	0.006810	0.000000	0.001719
PARIS 1ER ARRT	0.005391	0.000000	0.001529
PARIS 9E ARRT	0.004582	0.000000	0.000695
PARIS 3E ARRT	0.004272	0.000000	0.001394
PARIS 2E ARRT	0.002338	0.000000	0.000180
HAUTS-DE-SEINE	0.000000	0.026472	0.000000
SEINE-SAINT-DENIS	0.000000	0.057780	0.000000
VAL-DE-MARNE	0.000000	0.037694	0.000000

4. Analyse géographique

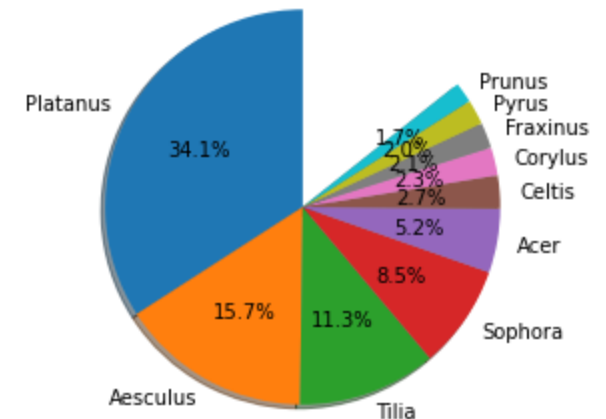
c) par genre d'arbres

- Le **Platane** est de loin l'arbre parisien le plus représenté avec **21%**
- Les **5 premiers genres** d'arbres sur environ 200 représente **60%** des arbres parisiens:
- Les **10 premiers 75%**

distribution des genres d'arbre à Paris



distribution des genres d'arbre sur les alignements



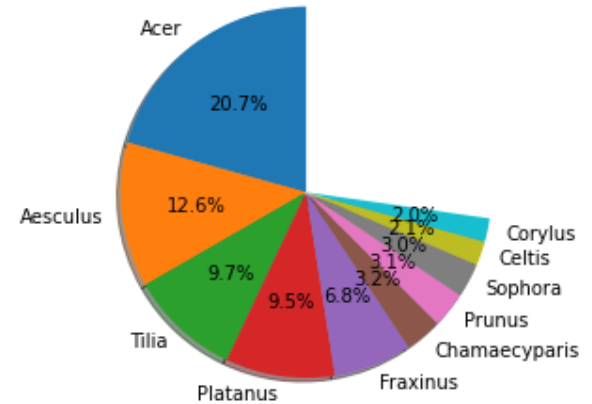
- Focus sur les alignements**
 - Le **platane avec 34%** est l'arbre qui règne en maître sur les alignements
 - Les 5 genres d'arbres les plus communs de Paris représente cette fois ci 75% des alignements

4. Analyse géographique

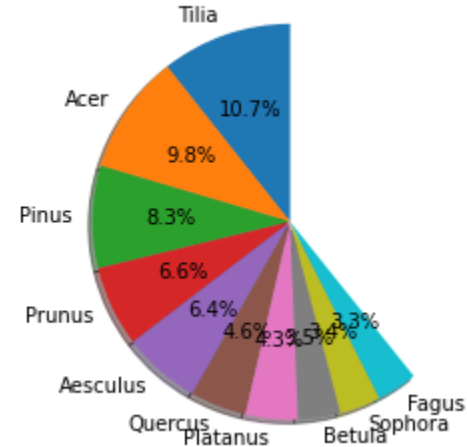
c) par genre d'arbres

- Focus sur les cimetières
 - Les **acers (érables)** y représentent 21% des arbres
 - Les platanes bien que représentés dans le top 5 ne constituent que 9.5% des arbres des cimetières
- Focus sur les jardins
 - La représentation des arbres y est un peu plus hétérogène qu'ailleurs:**
 - Les 5 premiers genres y représentent seulement 41% des arbres, Les 10 premiers 60%
 - Les platanes n'y représentent que 4.3% et sont à la 7^{ème} positions des genres les plus représentés

distribution des genres d'arbre dans les cimetières parisiens



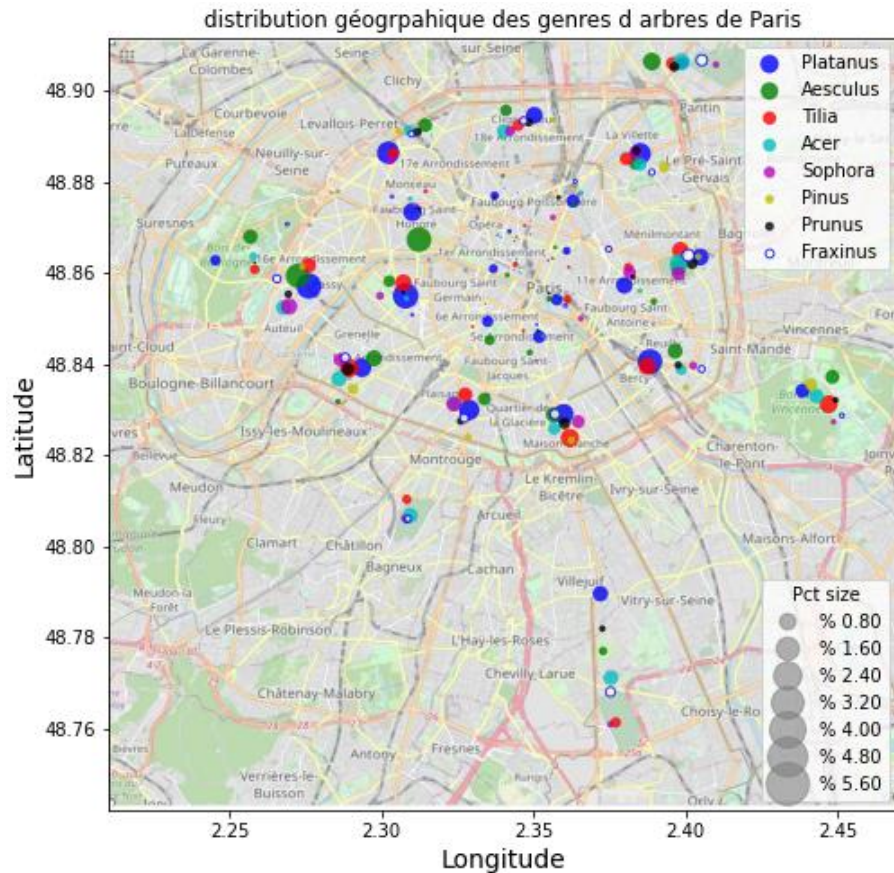
distribution des genres d'arbre dans les jardins parisiens



4. Analyse géographique

c) par genre d'arbres

- Les 5 genres d'arbres les plus représentés le sont également dans les arrondissements périphériques moitié sud de Paris: 16/15/14/13/12/20ème et bois de Vincennes.
- La plus forte représentation de **Platanus** se trouvent dans le **7ème et 16ème** et que celle d'**aesculus (marronniers)** se trouve dans le **8ème et 16ème**



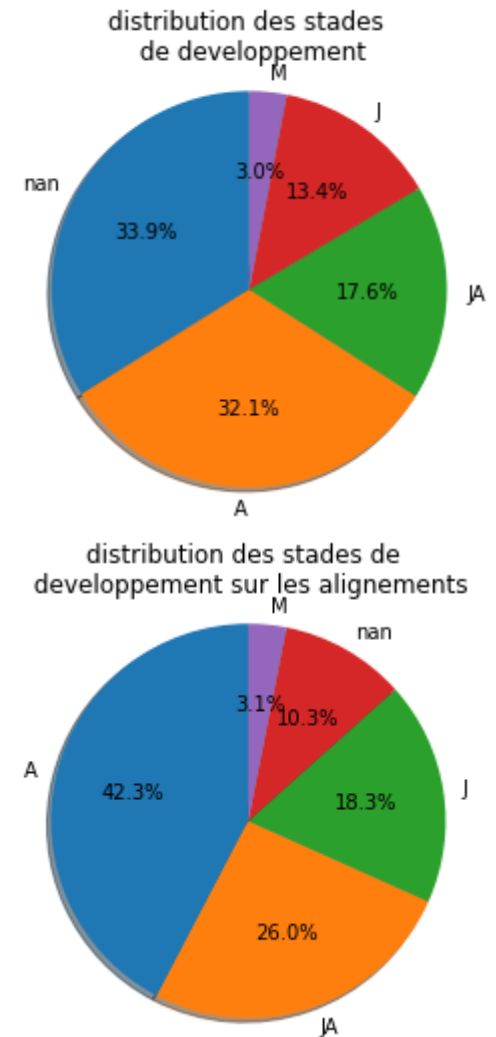
"Pourcentage d'arbres par genre et par arrondissement"

genre	Platanus	Aesculus	Tilia	Acer	Sophora
arrondissement					
PARIS 7E ARRD	%2.04	%0.43	%0.86	%0.09	%0.19
PARIS 16E ARRD	%2.00	%1.78	%0.68	%0.69	%0.81
PARIS 12E ARRD	%1.88	%0.70	%0.94	%0.39	%0.20
PARIS 17E ARRD	%1.55	%0.59	%0.38	%0.36	%0.22
PARIS 19E ARRD	%1.41	%0.48	%0.53	%0.55	%0.50
PARIS 14E ARRD	%1.30	%0.49	%0.66	%0.62	%0.70
PARIS 13E ARRD	%1.23	%0.75	%1.04	%0.48	%0.53
PARIS 15E ARRD	%1.10	%0.83	%1.10	%0.73	%0.53
PARIS 8E ARRD	%1.05	%1.91	%0.06	%0.08	%0.02
PARIS 20E ARRD	%0.89	%0.70	%0.87	%1.07	%0.57
PARIS 18E ARRD	%0.88	%0.42	%0.41	%0.62	%0.31
PARIS 11E ARRD	%0.84	%0.15	%0.27	%0.09	%0.50
VAL-DE-MARNE	%0.74	%0.22	%0.34	%0.68	%0.14
BOIS DE VINCENNES	%0.63	%0.62	%1.01	%0.68	%0.10
PARIS 10E ARRD	%0.55	%0.07	%0.04	%0.11	%0.11
PARIS 5E ARRD	%0.49	%0.12	%0.04	%0.04	%0.02
PARIS 4E ARRD	%0.42	%0.11	%0.20	%0.03	%0.09
SEINE-SAINT-DENIS	%0.41	%0.98	%0.51	%0.92	%0.15
PARIS 6E ARRD	%0.39	%0.34	%0.04	%0.00	%0.00
BOIS DE BOULOGNE	%0.35	%0.66	%0.29	%0.22	%0.00
HAUTS-DE-SEINE	%0.30	%0.11	%0.29	%0.69	%0.10
PARIS 1ER ARRD	%0.26	%0.08	%0.10	%0.03	%0.00
PARIS 9E ARRD	%0.23	%0.04	%0.01	%0.02	%0.04
PARIS 3E ARRD	%0.20	%0.06	%0.04	%0.01	%0.02
PARIS 2E ARRD	%0.07	%0.01	%0.02	%0.02	%0.04

4. Analyse géographique

d) par stade de developpement

- Avant traitement, 2/3 de la base possédaient un champs stade de développement renseigné.
- Sur ces 2/3, une bonne moitié est constituée d'arbres adultes.**
Très peu sont matures . Une autre bonne moitié peut être qualifiée de jeune/jeune adulte
- L'estimation des stades manquants ne change pas énormément la donne sur la répartition relative des stades
- Focus sur les alignements:
 - Seulement 10% des arbres sans stade renseigné**
 - Meilleur suivi électronique des alignements ,
comparativement aux autres domanialités
 - Beaucoup d'adultes, peu de matures
 - 18% des arbres sont strictement jeunes.**
 - Travail énorme
 - 19000 arbres plantés récemment.
 - forte proportion forcément abattue auparavant.
 - Les stades manquants des alignements pouvant être estimés sont en majorité des arbres jeunes

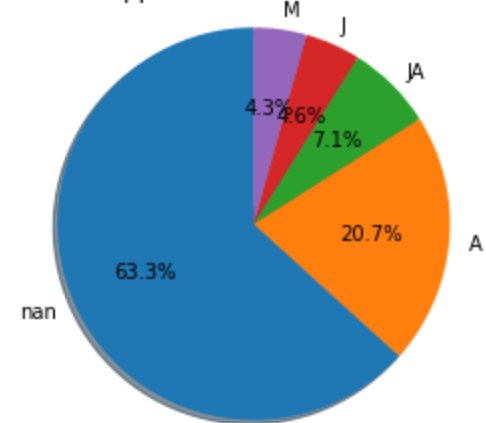


4. Analyse géographique

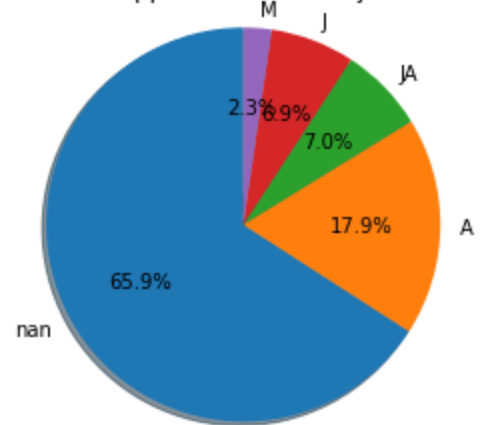
d) par stade de developpement

- Focus sur les Cimetières et les Jardins
 - Avant traitement , seulement 1/3 des arbres des cimetières et des jardins sont renseignés. **Ces domanialités sont donc clairement moins suivies** , du moins électroniquement parlant, que les alignements
 - Après estimation , une bonne moitié des arbres des cimetières ne possèdent toujours pas de stade de développement.
 - Le résultat est bien meilleur dans les jardins

distribution des stades de developpement dans les cimetières



distribution des stades de developpement dans les jardins

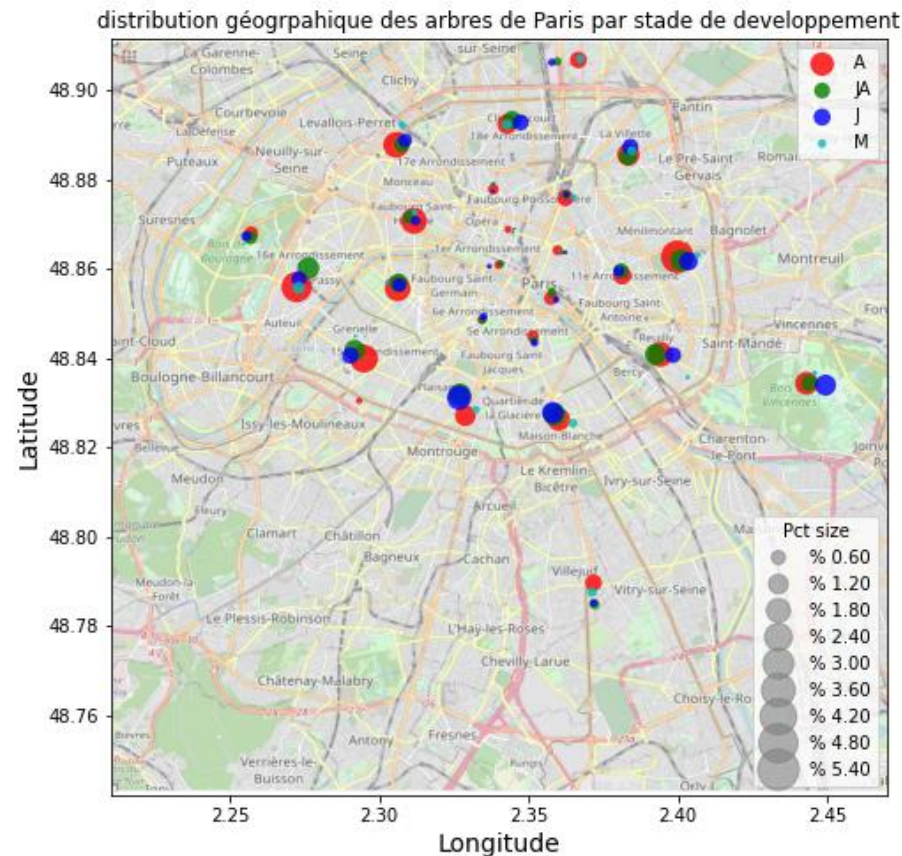


4. Analyse géographique

d) par stade de développement - avant estimation des stades manquants

- Avant estimation des stades manquants:
 - Le stade de développement **le plus représenté est l'adulte** (en majorité dans les 20^e/16^e/15^e)
 - Les **jeunes arbres** sont en majorité dans le 13/14^e
- et bois de Vincennes**

Pourcentage d'arbres par stade de développement et par arrondissement



stade_developpement	J	JA	A	M
arrondissement				
PARIS 20E ARRD	%1.11	%1.49	%3.46	%0.15
PARIS 16E ARRD	%0.74	%1.53	%3.01	%0.48
PARIS 15E ARRD	%0.86	%1.42	%2.52	%0.05
PARIS 8E ARRD	%0.29	%0.54	%2.19	%0.13
PARIS 7E ARRD	%0.60	%1.10	%2.14	%0.06
PARIS 17E ARRD	%0.62	%0.80	%2.13	%0.16
PARIS 12E ARRD	%0.80	%1.45	%2.04	%0.07
BOIS DE VINCENNES	%1.43	%0.82	%1.69	%0.10
PARIS 13E ARRD	%1.47	%1.61	%1.68	%0.25
PARIS 19E ARRD	%0.84	%0.95	%1.61	%0.29
PARIS 14E ARRD	%1.89	%1.60	%1.38	%0.20
PARIS 18E ARRD	%0.83	%1.06	%1.29	%0.33
PARIS 11E ARRD	%0.40	%0.75	%1.08	%0.01
VAL-DE-MARNE	%0.18	%0.33	%0.91	%0.30
SEINE-SAINT-DENIS	%0.18	%0.23	%0.91	%0.24
PARIS 10E ARRD	%0.13	%0.25	%0.88	%0.13
BOIS DE BOULOGNE	%0.33	%0.57	%0.63	%0.01
PARIS 4E ARRD	%0.14	%0.26	%0.59	%0.01
PARIS 5E ARRD	%0.18	%0.22	%0.44	%0.02
PARIS 9E ARRD	%0.05	%0.06	%0.38	%0.01
PARIS 3E ARRD	%0.05	%0.07	%0.33	%0.02
PARIS 6E ARRD	%0.18	%0.25	%0.26	%0.02
PARIS 1ER ARRD	%0.09	%0.16	%0.25	%0.00
PARIS 2E ARRD	%0.02	%0.03	%0.17	%0.00
HAUTS-DE-SEINE	%0.00	%0.01	%0.13	%0.01

4. Analyse géographique

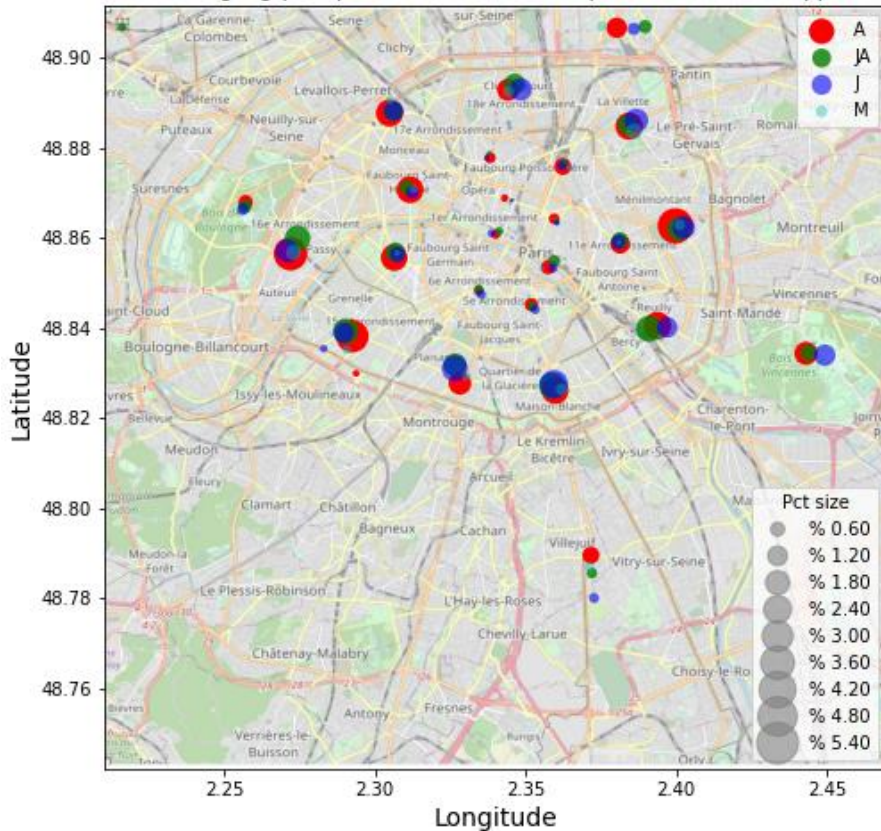
d) par stade de développement - après estimation des stades manquants

- Après estimation des stades manquants:
 - on a relativement la même analyse qu'au slide précédent.
 - Les autres arrondissements périphériques reçoivent une proportion un peu plus importante d'arbres jeunes**

Pourcentage d'arbres par stade de développement et par arrondissement

stade_developpement	J	JA	A	M
arrondissement				
PARIS 20E ARRD	%1.61	%1.73	%3.93	%0.37
PARIS 16E ARRD	%1.71	%2.09	%3.54	%0.72
PARIS 15E ARRD	%1.31	%2.05	%3.35	%0.25
PARIS 12E ARRD	%1.34	%2.15	%2.41	%0.18
PARIS 19E ARRD	%1.79	%1.40	%2.36	%0.85
PARIS 17E ARRD	%1.21	%1.14	%2.34	%0.23
PARIS 13E ARRD	%2.54	%2.23	%2.29	%0.47
PARIS 8E ARRD	%0.50	%0.61	%2.27	%0.21
PARIS 7E ARRD	%0.75	%1.16	%2.23	%0.10
BOIS DE VINCENNES	%1.44	%0.82	%1.69	%0.10
PARIS 14E ARRD	%2.10	%1.71	%1.61	%0.26
PARIS 18E ARRD	%1.58	%1.35	%1.52	%0.39
SEINE-SAINT-DENIS	%0.48	%0.56	%1.32	%0.31
PARIS 11E ARRD	%0.63	%0.87	%1.23	%0.06
VAL-DE-MARNE	%0.31	%0.36	%0.93	%0.32
PARIS 10E ARRD	%0.25	%0.26	%0.91	%0.15
PARIS 4E ARRD	%0.26	%0.39	%0.64	%0.06
BOIS DE BOULOGNE	%0.43	%0.57	%0.63	%0.01
PARIS 5E ARRD	%0.28	%0.28	%0.55	%0.05
PARIS 9E ARRD	%0.10	%0.07	%0.39	%0.02
PARIS 3E ARRD	%0.10	%0.10	%0.37	%0.03
PARIS 6E ARRD	%0.27	%0.26	%0.29	%0.05
PARIS 1ER ARRD	%0.20	%0.22	%0.27	%0.01
PARIS 2E ARRD	%0.05	%0.04	%0.17	%0.01
HAUTS-DE-SEINE	%0.18	%0.24	%0.14	%0.03

distribution géographique des arbres de Paris par stade de développement



4. Analyse géographique

d) par stade de developpement - par domanialité

- Les alignements des arrondissements périphériques mènent le jeu pour tous les stades.
- Dans le 93/94, les cimetières font de même
- Les arbres matures sont présents dans les alignements et jardins du 13/16 et 19^e et les cimetières extérieurs

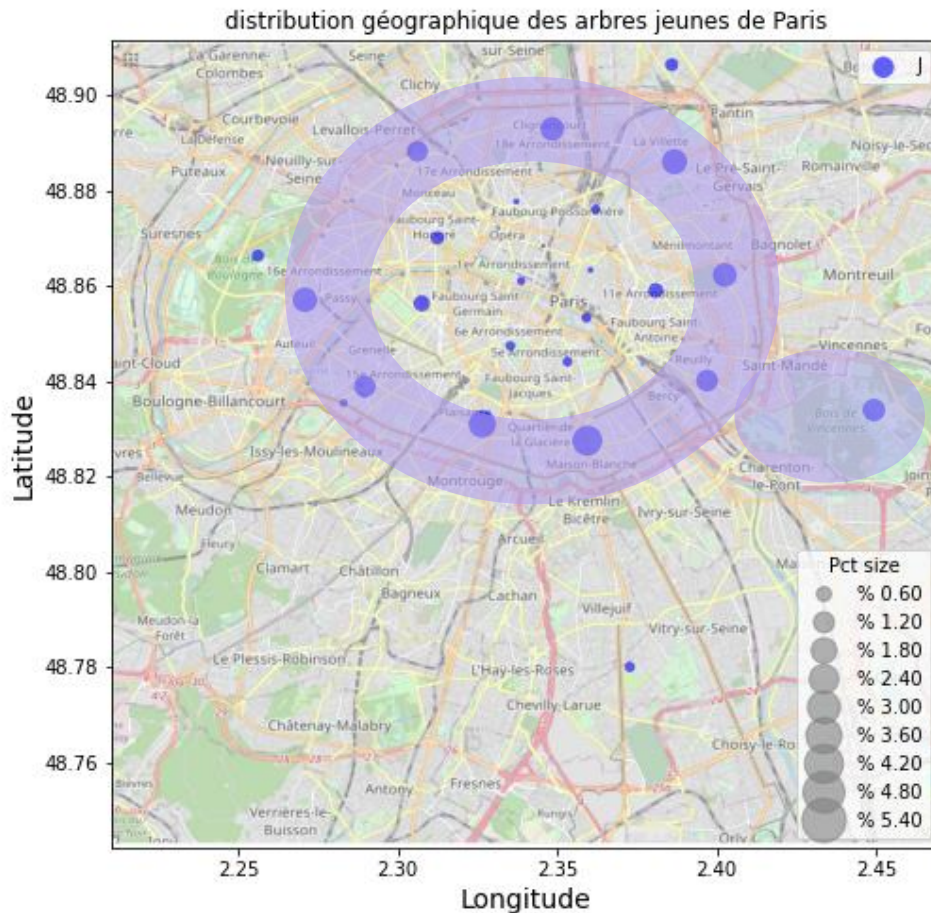
Nombre d'arbres pardomanialite et stade
de developpement et par arrondissement
avant estimation des stades manquants

domanialite stade_developpement arrondissement	Alignement				CIMETIERE				Jardin			
	A	J	JA	M	A	J	JA	M	A	J	JA	M
BOIS DE BOULOGNE	1,267	667	1,137	13	0	0	0	0	0	0	0	1
BOIS DE VINCENNES	1,464	1,889	1,312	28	0	0	0	0	1,913	970	337	169
HAUTS-DE-SEINE	0	0	0	0	266	9	23	12	0	0	0	0
PARIS 10E ARRD	1,668	213	386	177	0	0	0	0	11	0	0	27
PARIS 11E ARRD	1,900	615	1,307	4	0	0	0	0	106	25	73	2
PARIS 12E ARRD	3,312	1,245	2,582	52	5	0	2	0	355	81	100	58
PARIS 13E ARRD	2,328	2,425	2,719	436	0	0	0	0	81	123	73	39
PARIS 14E ARRD	1,211	2,925	2,092	196	151	254	469	0	1,068	464	541	159
PARIS 15E ARRD	4,226	1,348	2,484	46	0	2	0	0	346	70	117	27
PARIS 16E ARRD	5,356	1,158	2,691	746	50	0	7	3	372	179	217	173
PARIS 17E ARRD	3,187	809	1,258	89	364	152	107	141	418	174	110	71
PARIS 18E ARRD	1,542	929	1,608	420	362	150	117	132	192	250	158	54
PARIS 19E ARRD	2,660	1,178	1,664	535	0	0	0	0	77	56	29	36
PARIS 1ER ARRD	469	168	292	4	0	0	0	0	22	0	21	0
PARIS 20E ARRD	2,667	1,081	1,362	159	1,799	164	432	12	1,611	494	833	51
PARIS 2E ARRD	331	39	61	7	0	0	0	0	2	0	0	1
PARIS 3E ARRD	632	83	93	20	0	0	0	0	22	4	24	10
PARIS 4E ARRD	1,114	204	429	8	0	0	0	0	44	33	47	18
PARIS 5E ARRD	774	296	407	22	0	0	0	0	8	0	0	11
PARIS 6E ARRD	475	345	491	32	0	0	0	0	10	6	5	4
PARIS 7E ARRD	2,930	1,015	1,817	47	0	0	0	0	1,312	171	370	79
PARIS 8E ARRD	4,099	488	898	197	0	0	0	0	273	89	174	72
PARIS 9E ARRD	700	59	101	20	0	0	0	0	19	4	9	4
SEINE-SAINT-DENIS	0	0	0	0	1,813	367	465	473	0	0	0	0
VAL-DE-MARNE	0	0	0	0	1,802	380	651	607	0	0	0	0

Nombre d'arbres pardomanialite et stade
de developpement et par arrondissement
après estimation des stades manquants

domanialite stade_developpement arrondissement	Alignement				CIMETIERE				Jardin			
	A	J	JA	M	A	J	JA	M	A	J	JA	M
BOIS DE BOULOGNE	1,268	858	1,138	13	0	0	0	0	0	0	0	1
BOIS DE VINCENNES	1,464	1,907	1,313	28	0	0	0	0	1,914	972	337	169
HAUTS-DE-SEINE	0	0	0	0	273	370	473	59	0	0	0	0
PARIS 10E ARRD	1,676	343	388	180	0	0	0	0	54	66	26	74
PARIS 11E ARRD	1,952	914	1,327	28	0	0	0	0	347	130	289	77
PARIS 12E ARRD	3,334	1,691	2,608	69	5	0	2	0	857	511	1,277	210
PARIS 13E ARRD	2,570	3,660	3,016	537	0	0	0	0	1,007	735	914	362
PARIS 14E ARRD	1,385	3,212	2,229	212	343	301	519	64	1,116	513	565	195
PARIS 15E ARRD	4,254	1,484	2,530	53	1	17	0	0	1,900	654	1,280	421
PARIS 16E ARRD	5,379	1,931	2,704	766	65	34	121	5	957	675	676	577
PARIS 17E ARRD	3,209	1,115	1,290	97	366	154	109	141	763	828	646	179
PARIS 18E ARRD	1,544	1,256	1,613	426	368	200	124	133	435	995	369	161
PARIS 19E ARRD	2,827	1,916	1,791	623	13	29	10	42	1,316	989	745	980
PARIS 1ER ARRD	477	279	310	12	0	0	0	0	54	118	122	9
PARIS 20E ARRD	2,667	1,158	1,368	159	2,478	568	654	347	1,671	777	965	91
PARIS 2E ARRD	331	62	67	8	0	0	0	0	4	8	7	11
PARIS 3E ARRD	635	106	94	20	0	0	0	0	82	64	84	37
PARIS 4E ARRD	1,122	377	432	11	0	0	0	0	123	86	314	116
PARIS 5E ARRD	779	361	410	22	0	0	0	0	217	123	120	81
PARIS 6E ARRD	477	362	492	32	0	0	0	0	80	161	26	74
PARIS 7E ARRD	2,943	1,155	1,824	52	0	0	0	0	1,465	329	483	143
PARIS 8E ARRD	4,099	679	907	197	0	0	0	0	430	307	309	231
PARIS 9E ARRD	700	95	101	20	0	0	0	0	46	38	30	21
SEINE-SAINT-DENIS	0	0	0	0	2,641	967	1,120	612	0	0	0	0
VAL-DE-MARNE	0	0	0	0	1,833	627	710	631	0	0	0	0

5. Focus sur les jeunes arbres



- Les jeunes arbres: synonyme d'abattage, de remplacement d'arbres potentiellement malades, puis d'entretien
- **La majorité des arbres jeunes se trouvent dans les alignements des arrondissements de la périphérie de Paris.**
- **Les 13/14^e arrivent en tête.** Ensuite viennent le bois de Vincennes et les autres arrondissement périphériques (cf slide suivant)
- **Le platane est de loin l'arbre le plus représenté chez les jeunes.** Les marronniers viennent ensuite.
- (cf slide suivant)

5. Focus sur les jeunes arbres

nombre d arbres jeunes par arrondissement
avant et après estimation des stades manquants

	arrondissement	nb_arbres_av_est	nb_arbres_ap_est
0	PARIS 14E ARRD	3777	4205
1	PARIS 13E ARRD	2946	5076
2	BOIS DE VINCENNES	2859	2879
3	PARIS 20E ARRD	2220	3217
4	PARIS 15E ARRD	1726	2626
5	PARIS 19E ARRD	1678	3580
6	PARIS 18E ARRD	1654	3154
7	PARIS 12E ARRD	1605	2690
8	PARIS 16E ARRD	1488	3414
9	PARIS 17E ARRD	1239	2423
10	PARIS 7E ARRD	1207	1507
11	PARIS 11E ARRD	799	1257
12	BOIS DE BOULOGNE	667	858
13	PARIS 8E ARRD	584	1004
14	SEINE-SAINT-DENIS	367	967
15	PARIS 5E ARRD	362	555
16	PARIS 6E ARRD	361	533
17	VAL-DE-MARNE	360	627
18	PARIS 4E ARRD	281	528
19	PARIS 10E ARRD	261	507
20	PARIS 1ER ARRD	176	405
21	PARIS 3E ARRD	109	202
22	PARIS 9E ARRD	95	191
23	PARIS 2E ARRD	44	91
24	HAUTS-DE-SEINE	9	370

nombre d arbres jeunes par domanialite
avant et après estimation des stades manquants

	domanialite	nb_arbres_av_est	nb_arbres_ap_est
0	Alignement	19179	24921
1	Jardin	3193	9079
2	DASCO	1745	2447
3	CIMETIERE	1458	3267
4	DJS	766	913
5	DFPE	292	419
6	PERIPHERIQUE	241	1815

statistiques basiques des arbres jeunes
des 10 premiers genres représentés
après estimation des stades manquants

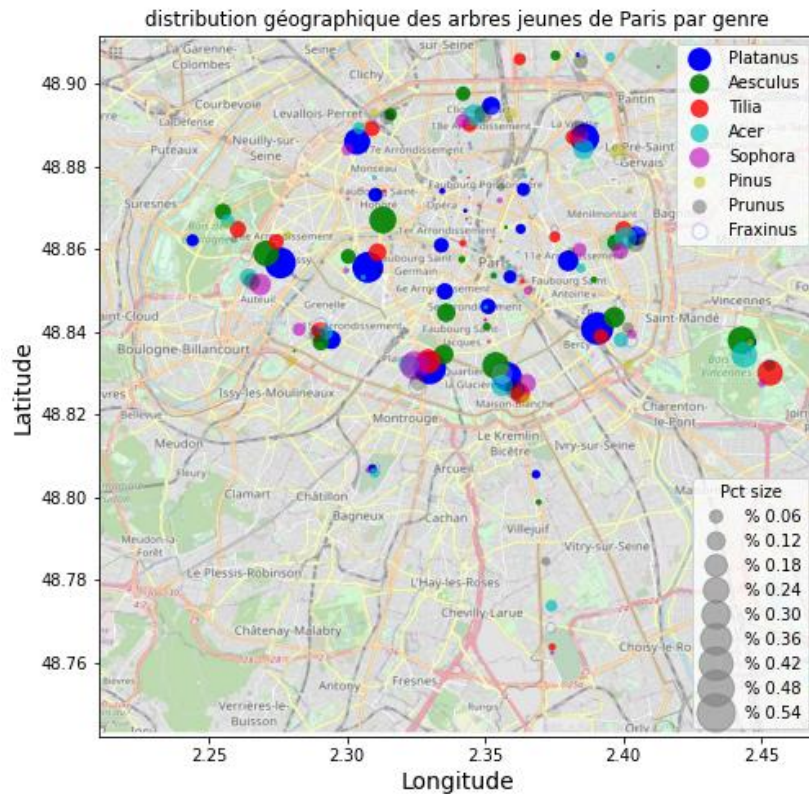
nombre d arbres jeunes par genre
avant et après estimation des stades manquants

	genre	nb_arbres_av_est	nb_arbres_ap_est
0	Platanus	5469	7353
1	Aesculus	3389	4468
2	Tilia	2228	3285
3	Acer	1977	3441
4	Sophora	1882	2524
5	Prunus	1386	2490
6	Pyrus	939	1760
7	Fraxinus	882	1510
8	Quercus	794	1241
9	Celtis	733	1058

	nb	hauteur moy	circ moy	hauteur med	circ med	hauteur min	circ min	hauteur max	circ max	hauteur std	circ std
genre											
Platanus	7,353.0	6.4	36.4	5.0	30.0	3.0	2.0	25.0	260.0	2.1	19.0
Aesculus	4,468.0	6.0	41.8	5.0	37.0	1.0	3.0	16.0	290.0	2.0	21.4
Acer	3,441.0	5.5	35.8	5.0	30.0	1.0	3.0	14.0	180.0	1.5	18.0
Tilia	3,285.0	6.0	40.7	5.0	35.0	2.0	2.0	20.0	173.0	2.1	21.9
Sophora	2,524.0	5.4	35.2	5.0	30.0	2.0	2.0	18.0	211.0	1.2	16.2
Prunus	2,490.0	4.3	30.0	5.0	25.0	1.0	4.0	8.0	181.0	1.3	13.8
Pyrus	1,760.0	5.2	28.0	5.0	25.0	1.0	3.0	10.0	91.0	0.9	11.2
Fraxinus	1,510.0	5.2	34.8	5.0	30.0	3.0	10.0	18.0	108.0	1.2	13.7
Quercus	1,241.0	5.5	31.9	5.0	30.0	1.0	6.0	15.0	110.0	1.7	15.1
Celtis	1,058.0	5.0	34.6	5.0	26.0	2.0	15.0	10.0	120.0	1.1	15.9

5. Focus sur les jeunes arbres

Nombre d'arbres jeune par genre
et par arrondissement



genre	Platanus	Aesculus	Tilia	Acer	Sophora	Pinus	Prunus	Fraxinus
arrondissement								
PARIS 14E ARRD	804	292	446	316	617	41	221	214
PARIS 12E ARRD	779	325	167	167	70	77	95	104
PARIS 16E ARRD	725	495	176	236	350	61	214	87
PARIS 7E ARRD	718	165	237	22	28	1	29	1
PARIS 13E ARRD	637	546	286	344	262	89	298	346
PARIS 19E ARRD	632	103	160	312	234	214	207	88
PARIS 17E ARRD	518	118	179	108	105	46	162	89
PARIS 11E ARRD	327	32	96	70	144	5	46	38
PARIS 15E ARRD	267	196	194	145	129	85	210	110
PARIS 20E ARRD	265	193	211	370	207	54	222	136
PARIS 18E ARRD	252	158	172	347	153	117	232	134
PARIS 6E ARRD	207	261	6	2	0	0	6	0
PARIS 5E ARRD	181	48	11	17	13	1	25	5
PARIS 1ER ARRD	167	36	41	13	0	0	3	2
PARIS 8E ARRD	154	553	12	19	8	10	27	6
PARIS 10E ARRD	130	3	5	18	28	0	53	3
PARIS 4E ARRD	129	32	18	17	60	1	26	16
BOIS DE BOULOGNE	110	202	200	142	1	48	43	0
BOIS DE VINCENNES	94	583	492	502	47	35	110	13
PARIS 3E ARRD	78	12	1	6	6	0	10	0
HAUTS-DE-SEINE	63	7	12	75	21	1	0	17
VAL-DE-MARNE	58	27	39	104	14	0	59	66
PARIS 9E ARRD	28	4	6	12	5	1	32	0
SEINE-SAINT-DENIS	16	76	112	72	9	0	155	35
PARIS 2E ARRD	14	1	6	5	13	0	5	0



CONCLUSIONS

1. Les Conclusions

- La majorité des arbres se trouvent sur **les arrondissements périphériques** de Paris (le 15^{ème} arrivant en tête.
 - On peut y ajouter le bois de Vincennes , le 7^è et le 8^è et enfin les cimetières du 93 et 94
- Les **alignements** concentrent la majorité des arbres parisiens. Les Jardins et les cimetières viennent ensuite.
- Le **platane** est de loin l'arbre le plus courant sur Paris. Les marronniers, tilleuls, érables et sophoras suivent pour représenter à eux 5: 60% des arbres.
 - La plus forte concentration de Platane se trouve par ordre décroissant dans le 7^e/16^e/12^e
 - La plus forte concentration de Marronniers se trouve par ordre décroissant dans le 8^e/16^e
- Le **stade de développement le plus représenté est l'adulte** (en majorité dans les 20^e/16^e/15^e) au contraire des matures.
- **Les jeunes arbres qui nécessitent beaucoup de travail (abattage/entretien) sont bien représentés dans les alignements du 13^e/14^e arrondissements** et les autres arrondissement périphériques, bois de Vincennes compris
 - Les platanes et les marronniers sont les jeunes arbres les plus courants
- **Conseil:**
 - 1) Améliorer la base:
 - *en étant plus propre*
 - *en renseignant plus de champs permettant une analyse plus profonde*
 - 2) Optimisation des ressources: D'un point de vue analytique, les alignements des arrondissements périphériques avec leurs platanes et marronniers doivent concentrer la majorité des efforts.
 - *Ces arbres rencontrent des maladies bien connues et leurs alignements favorisent la propagation des épidémies. Il faut en abattre certains , replanter, puis veiller comme des parents ces nouveaux nés.*
 - *Les jardins, jamais laissés à l'abandon, nécessitent également de lourds entretiens mais moins visibles dans les chiffres de notre base*
- Pour aller plus loin:
 - Possibilité de modéliser des anomalies de croissance