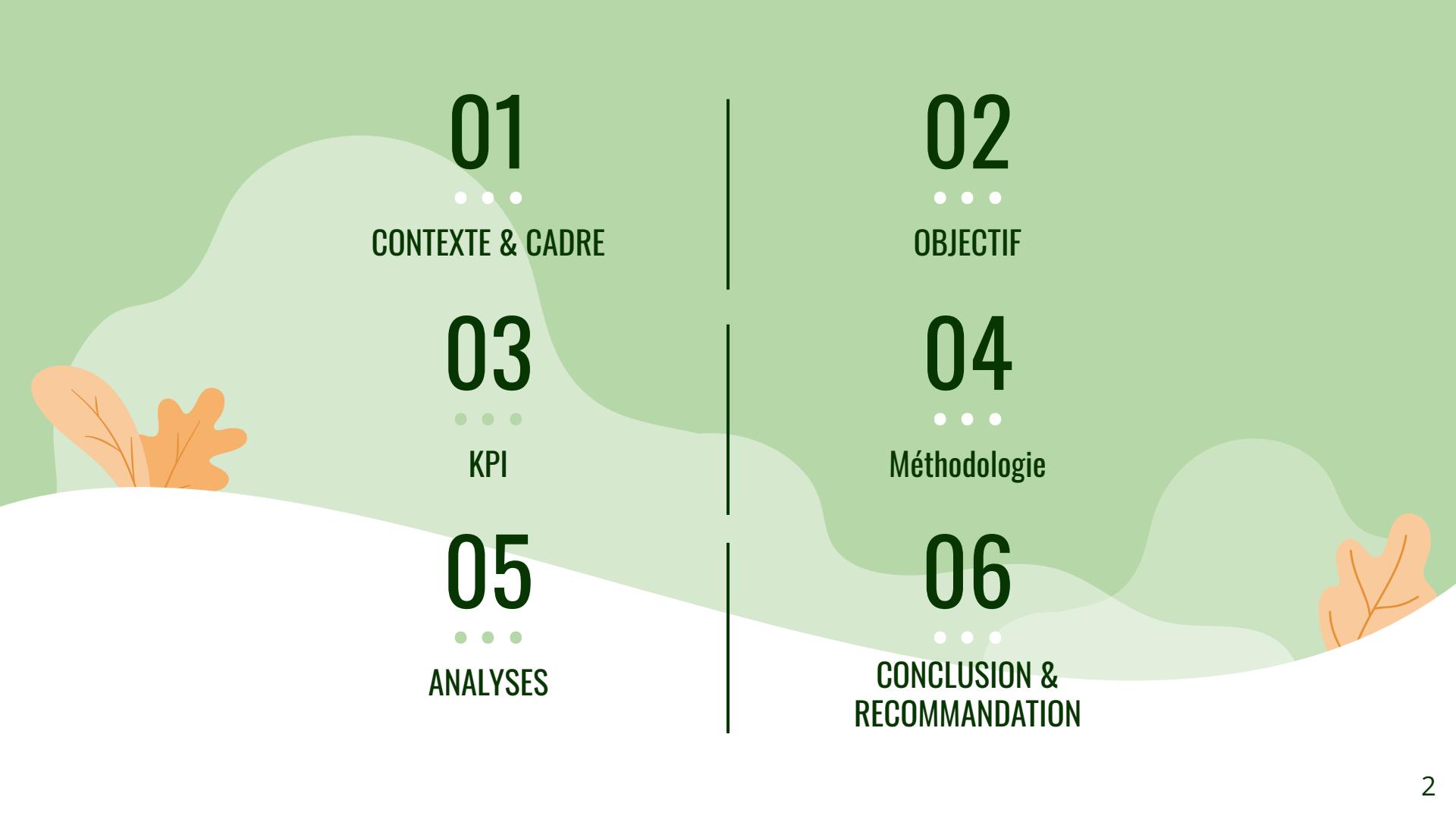


Capstone III

Vélib'

Nassriat Youssouf, Kamala Sababady, Tharsana Balachandran
Groupe 4





01

...
CONTEXTE & CADRE

03

...
KPI

05

...
ANALYSES

02

...
OBJECTIF

04

...
Méthodologie

06

...
CONCLUSION &
RECOMMANDATION

01

...

CADRE & CONTEXTE





Velib'

Service de vélo partagé

But : Développer la mobilité dans la métropole.

Notre Rôle :

- Études des données
- Évaluation des activités de Vélib
- Stratégies d'améliorations



Cadre

Objectif

KPI

Méthodologie

Analyses

Conclusion

Cadre

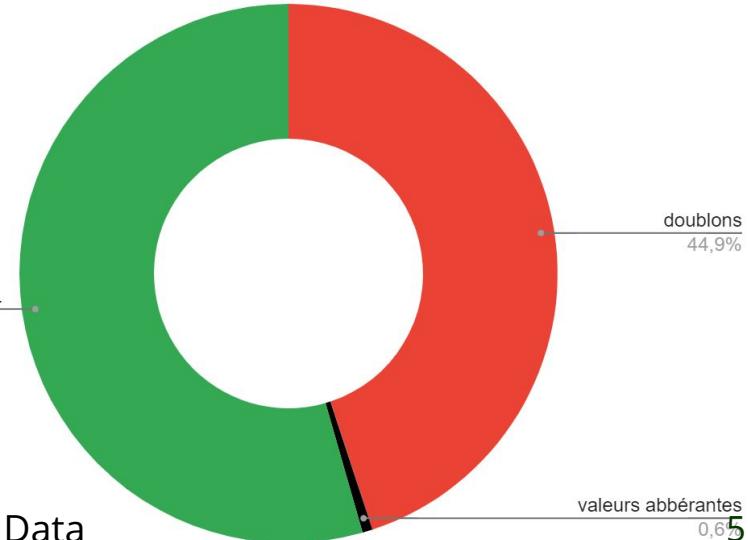
- 3 Data Analyst
- Analyse ~ 4 jours
- Données disponibles : mars 2022, avril 2022 et mai 2022
- Période : avril 2022
(4 jours de données manquantes)
- Données de 1414 stations disponibles sur 60 arrondissements en Île-de-France

Santé de la Data

Données exploita...
54,4%

doublons
44,9%

valeurs abbréantes
0,65



02 ... OBJECTIF

Comment satisfaire les clients ?

- Quel est l'activité pour chaque station ?
- Comment varie l'activité en fonction des grandes gares ?



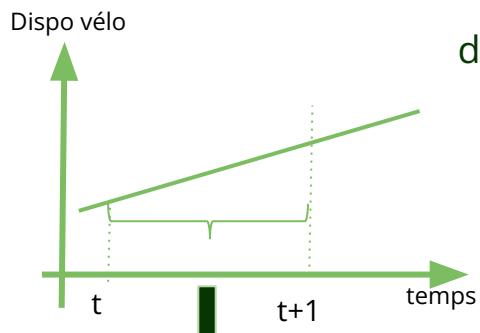
03

KPI





Quels KPI ?



Disponibilité des vélos

Nombre de vélos disponibles / Capacité de la station

Activité des stations par jour

Différence des activités vélo par station

Seuil minimal de disponibilité vélos

1 Vélo / Capacité de la station

04

...

METHODOLOGIES





Data Préparation

Pandas sous Python

Chargement des données

- 1 fichier plat pour le mois d'avril

Nettoyage des données

- Suppression des valeurs aberrantes
- Suppression des valeurs nulles
- Suppression des doublons

Enrichissement des données

- Disponibilité des vélos
- L'activité de la station
- Capacité réelle
- Catégorie de l'activité
- Catégorisation des capacités
- Calcul des distances entre stations et les grandes gares

Données externes

- Données de gares en Île-de-France
Localisation des gares sur Paris et ses banlieues proches

05

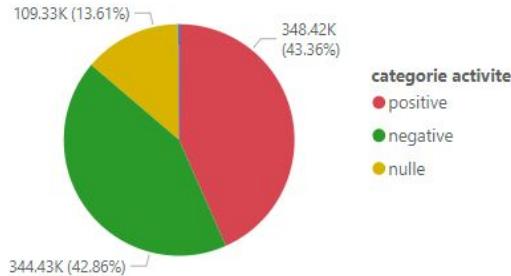
...

ANALYSES



Toutes les stations passent par les 3 phases d'activités

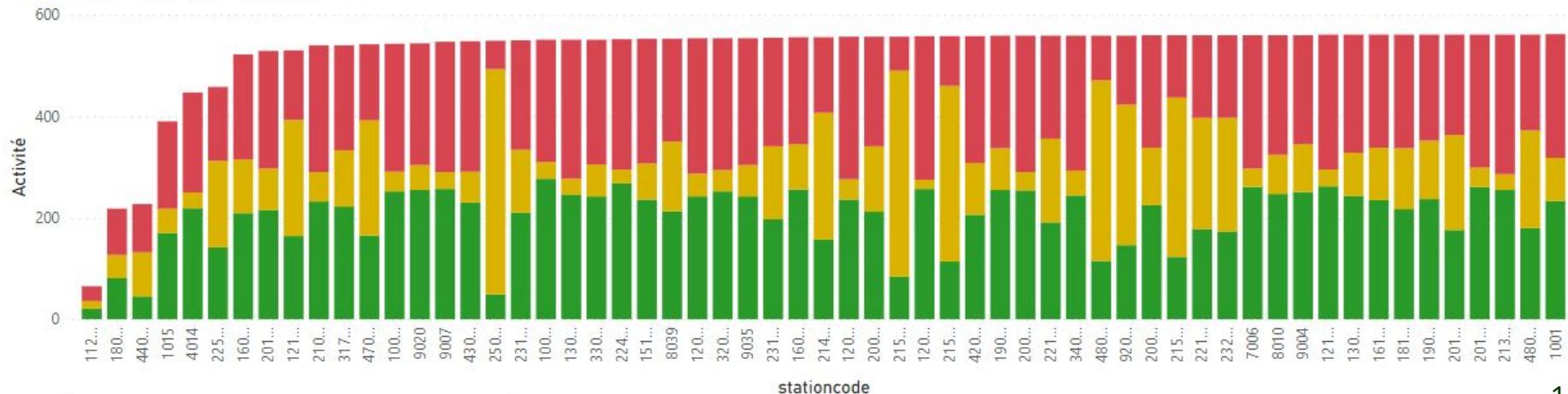
Nombre de stationcode par catégorie activité



Catégorie activités :

- Nulle (Jaune) ⇒ pas d'activité
- Positive (Rouge) ⇒ retour des vélos
- Négative (Vert) ⇒ sortie des vélos

catégorie activité ● negative ● nulle ● positive



Contexte

Objectif

KPI

Méthodologie

Analyses

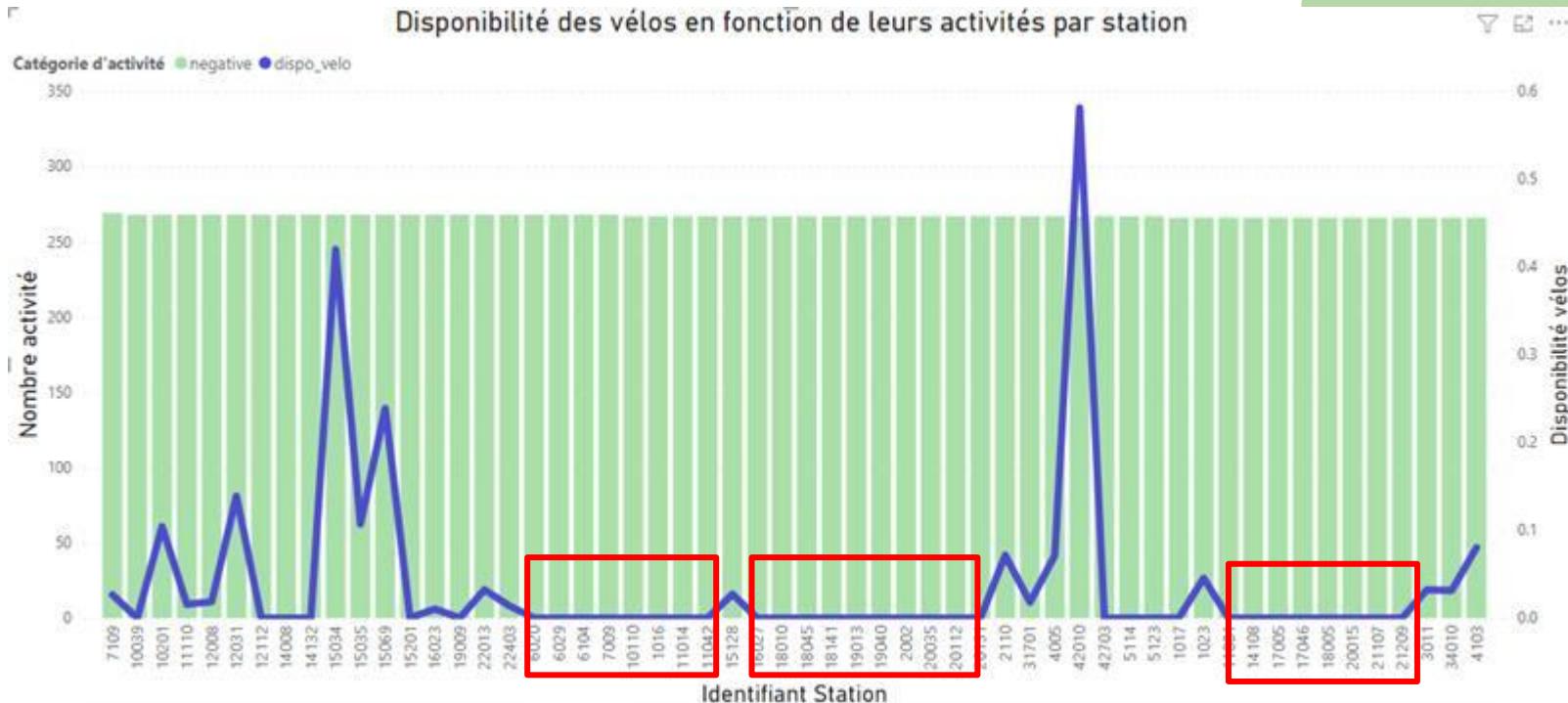
Conclusion



Condition Minimum :

Avoir au moins 1 vélo de
disponible par station vélib'

917 stations ont une disponibilité vélo nulle



- Forte activité négative
- Disponibilité des vélos minimum sur le mois

Contexte

Objectif

KPI

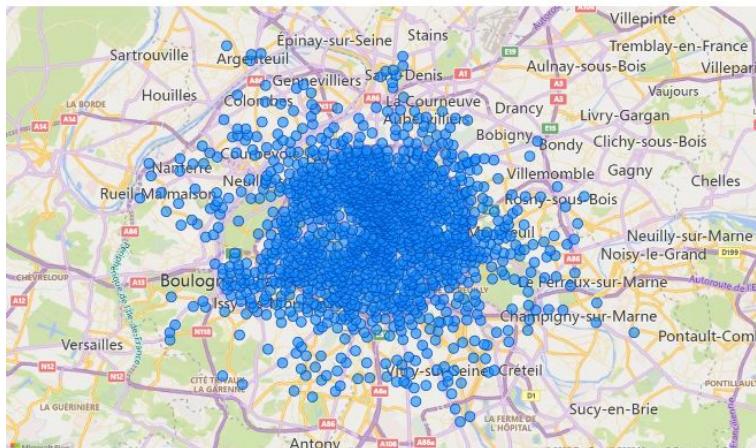
Méthodologie

Analyses

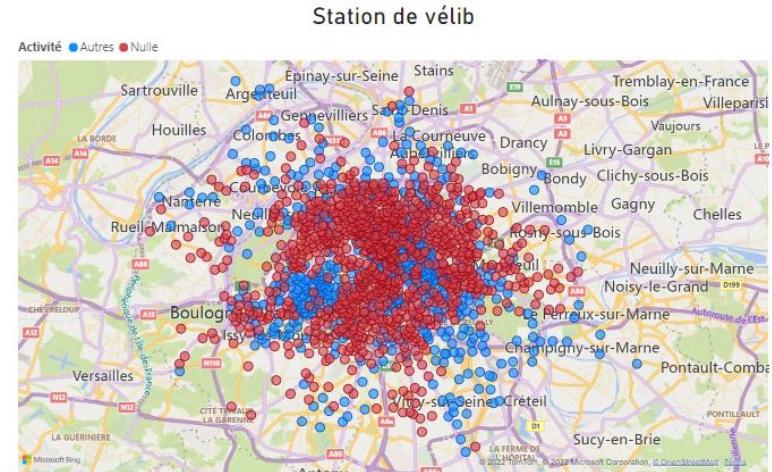
Conclusion

917 stations sur 1414 ont une disponibilité vélo nulle

Station de vélib



Station de vélib



1110 sur 1414 ne respectent pas la condition minimale



➤ Forte activité négative

➤ Disponibilité des vélos minimum sur le mois
Seuil_vélo = 1 / Capacité de la station

Disponibilité vélo < Seuil min ⇒ Non- Respect de la condition

Contexte

Objectif

KPI

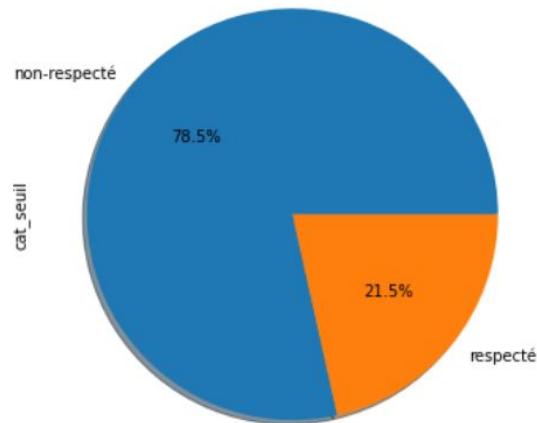
Méthodologie

Analyses

Conclusion

Manque de vélos pendant une activité forte

Respect de la condition minimum pour les station à forte activité



- Au moins une fois dans le mois, il y a 78.5% des stations qui ne respectent pas notre condition minimum.
- Manque de vélos □ ajout des vélos

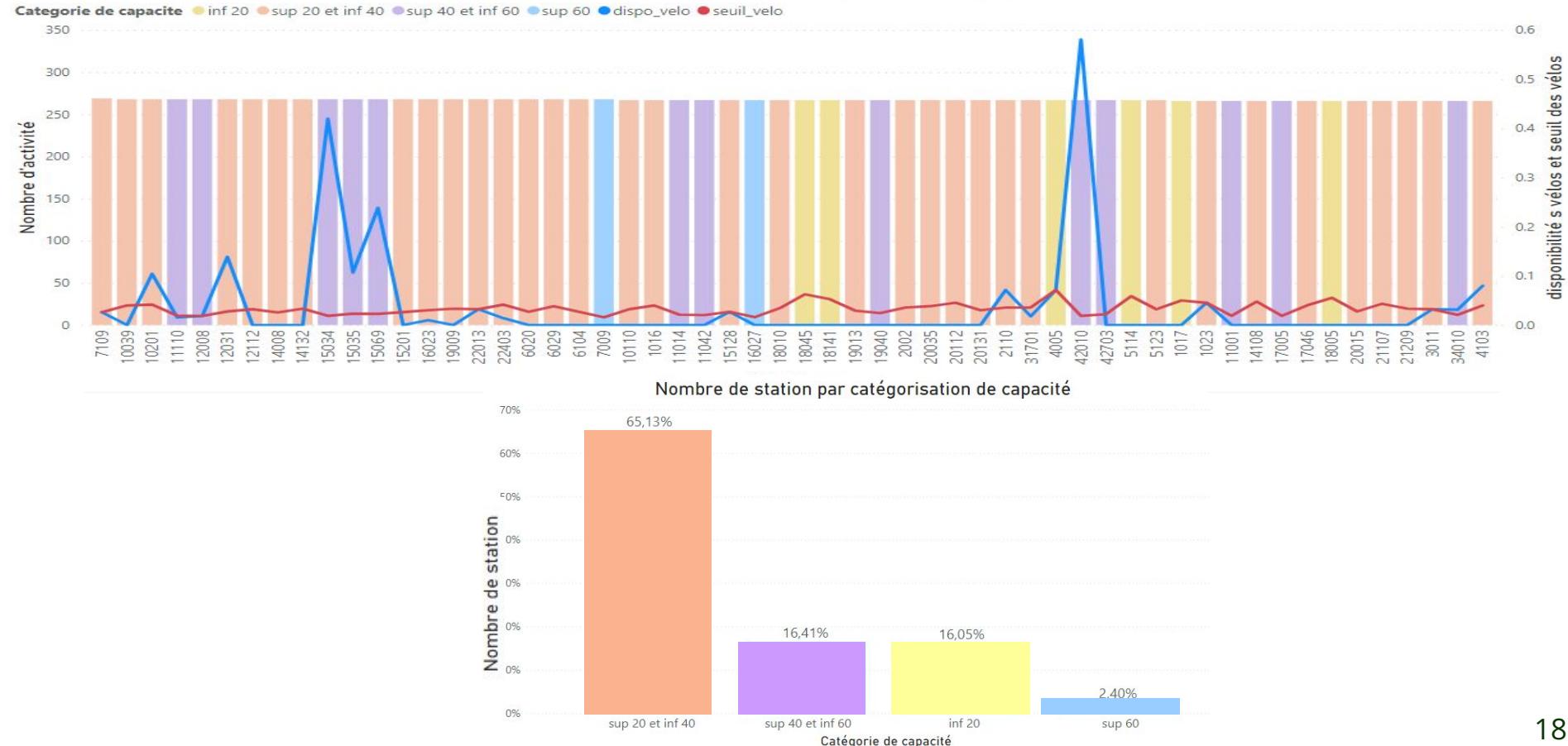


Attention :

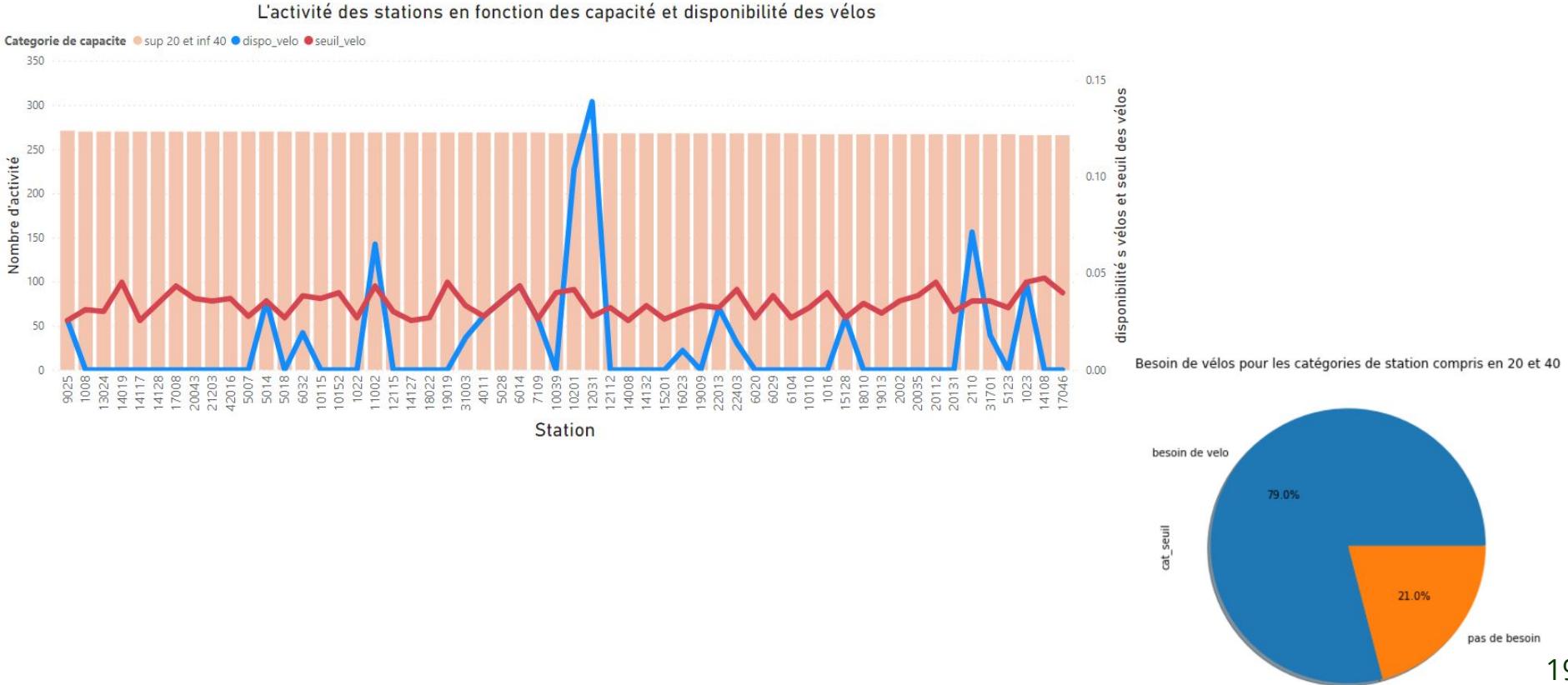
- Capacité de la station
- Modalité répartition des vélos

La majorité des stations ont une capacité comprise entre 20 et 40

L'activité des stations en fonction des capacité et disponibilité des vélos



79% des stations ont des besoins de vélo au moins 1 fois dans le mois



Contexte

Objectif

KPI

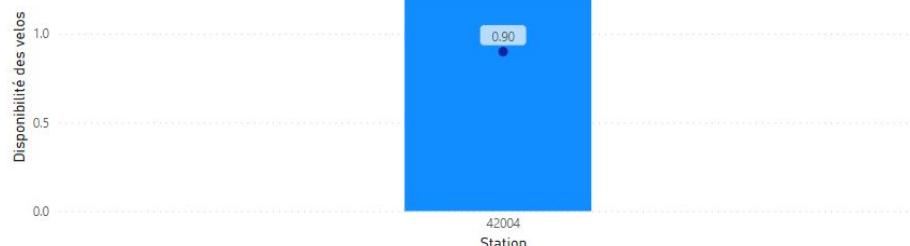
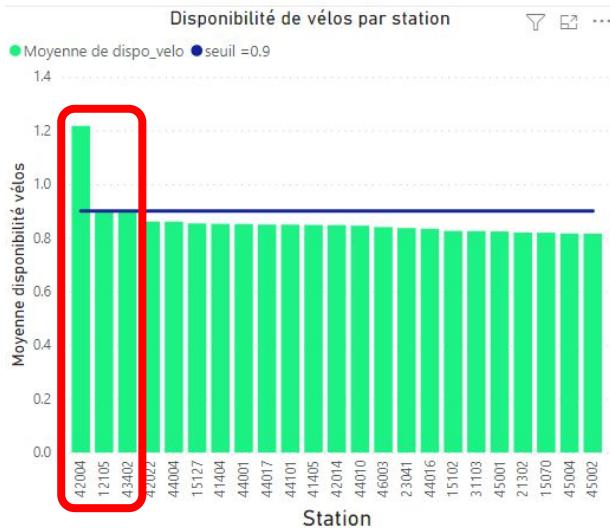
Méthodologie

Analyses

Conclusion

Dépassement des vélos disponibles pour la Westermeyer - Paul Vaillant-Couturier

Seuil défini sur le critère d'au moins 1 bornette disponible dans la station



stationcode gare à proximité

1140

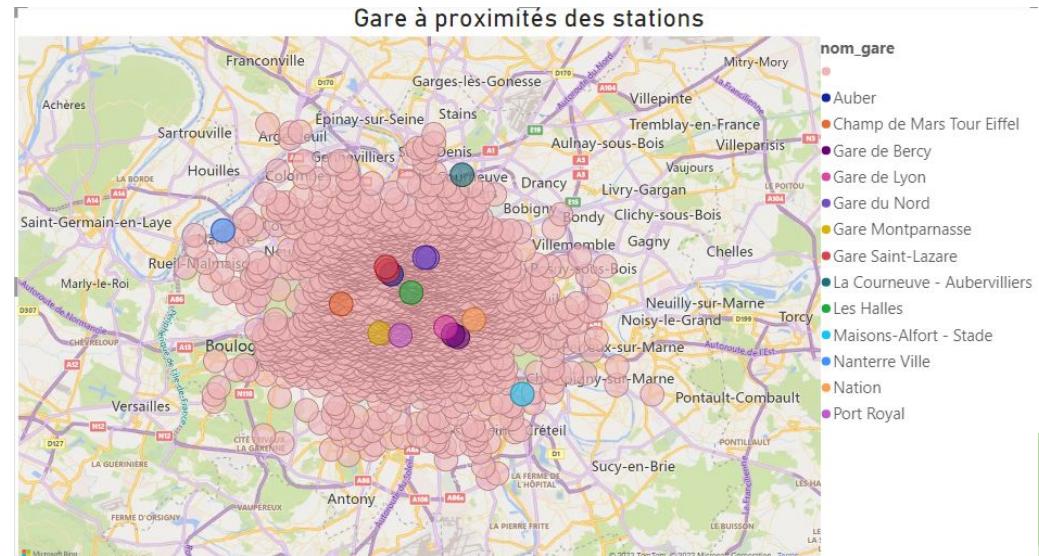
42004

Gare de Bercy

Répartition des stations en fonction des gares

Hypothèse :

Les gares influencent l'activité des stations Vélib



Contexte

Objectif

KPI

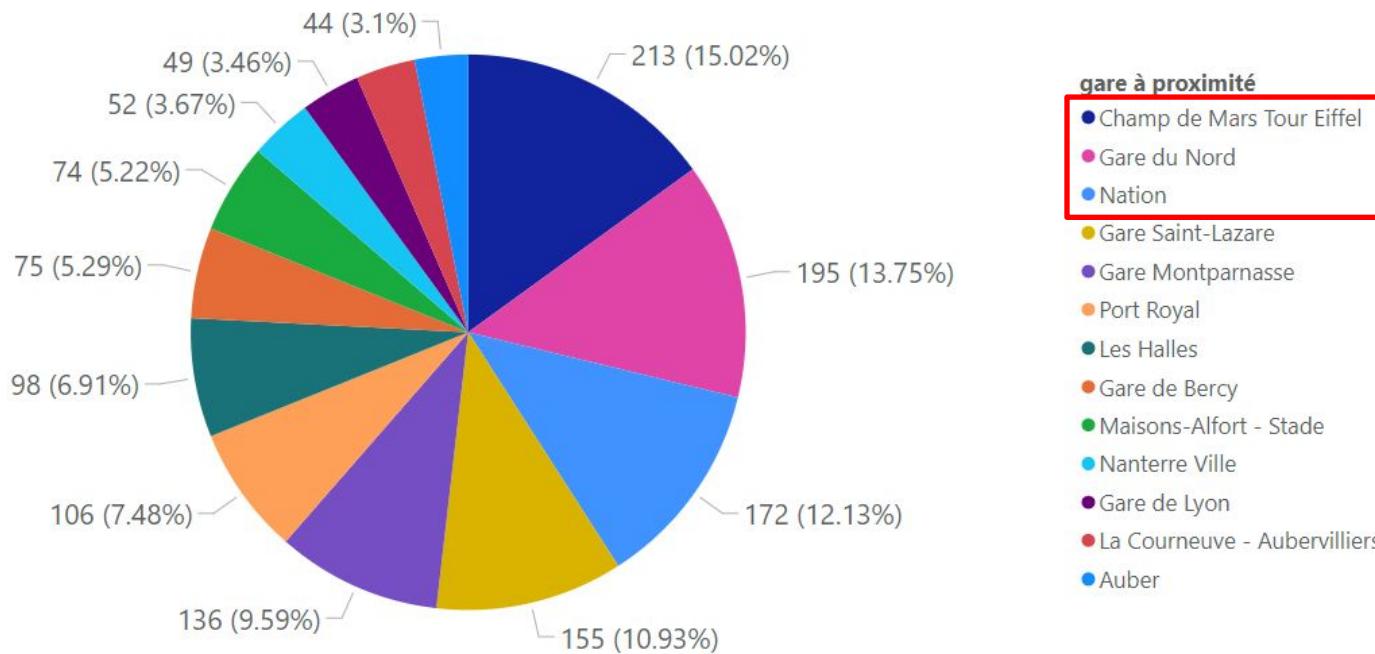
Méthodologie

Analyses

Conclusion

Critères de sélection du choix des grandes gares

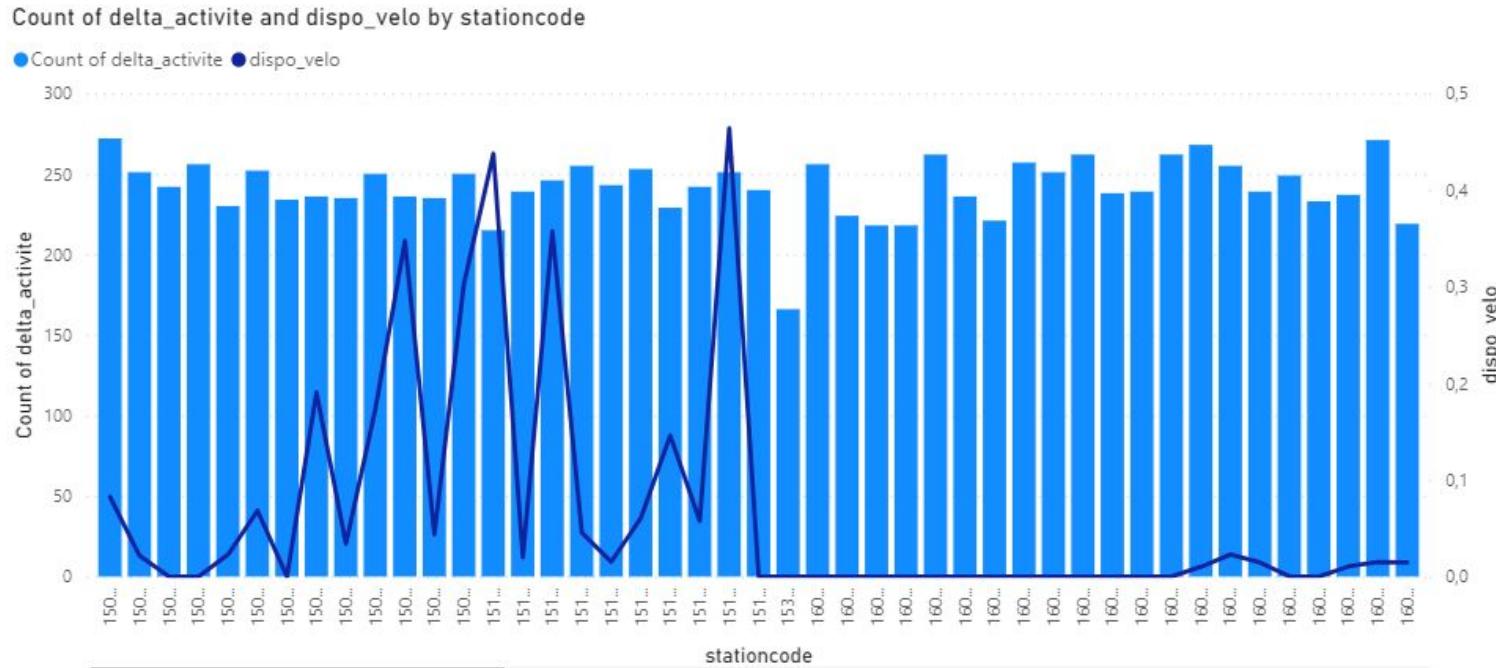
Nombre de station par gare à proximité



Gare : Champ de Mars Tour Eiffel

Catégorie activité : activité négative

Catégorie de capacité : supérieur à 20 et inférieur à 40



On remarque qu'il y a une mauvaise répartition de vélo sur différentes stations pour un créneau horaire données

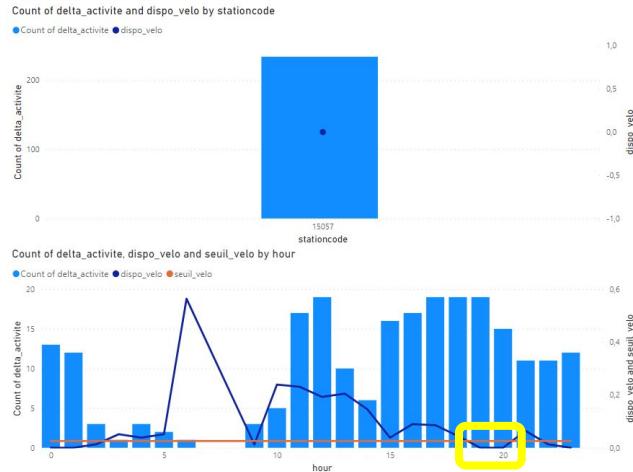
CAS PARTICULIER

Gare : Champ de Mars Tour Eiffel

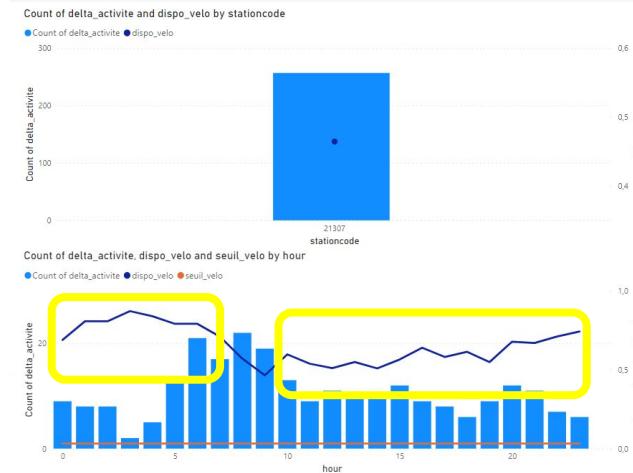
Catégorie activité : activité négative

Catégorie de capacité : supérieur à 20 et inférieur à 40

Cas de manque



Cas de dépassement



Bon indicateur \Rightarrow seuil_vélo < disponibilité vélo < 0.9

06

...

CONCLUSION



Contexte

Objectif

KPI

Méthodologie

Analyses

Conclusion



Conclusion pour le mois d'avril 2022

Analyse Business :

- Indisponibilité des vélos en cas de forte activité au moins 1 fois par mois
- Certaines stations ont un nombre de vélos qui peut dépasser la capacité de la station
- Stations placées stratégiquement autour des grandes gares
- Inégalité des répartitions des vélos



Axe d'amélioration

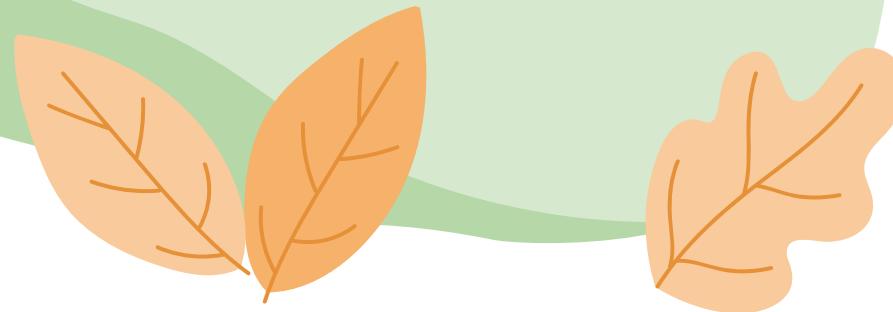
- Manque de données sur la gestion des ressources en internes (chargement des vélos)
- Problèmes de performances machines
- Plus de temps :
 - Permet d'avoir une meilleure généralisation de l'analyse
 - Donner des recommandations exactes
 - Généralisation pour toutes les gares d'Ile-de-France

Recommandation

- Gestion des ressources : Revoir la répartition des vélos par station
 - Rechargement de vélos 2 fois par jours
 - Camions pour un nombre de stations données en fonction des grandes gares à proximité
 - Camion ayant une capacité de 20 vélos
 - 2 employés qui seront responsables de cette tache par camion
- Ou Achat de nouveaux vélos pour combler cette inégalité



**Merci de votre
écoute**



Annexes

...



Préparation des données

Python pandas

- Choix de suppression des colonnes ping time ⇒ pour supprimer les lignes doublons par rapport “duedate”
- Suppression des doublons dû aux mauvaises dates de ping_time (hors avril)
- Suppression des capacités = 0 ⇒ car pas de données exploitables
- Suppression des Nan ⇒ stations pas encore installées
- Pour faciliter le temps de calcul ⇒ division du dataset en 3 sous dataset selon les stations (472, 473, 469)

Données externes :

<https://data.iledefrance-mobilites.fr/explore/dataset/emplacement-des-gares-idf-data-generalisee/table/>

Analyse des données

Python

- Création de colonne calculée pour l'analyse
- 2 Catégorisations
- Calculs de matrice de distance pour chaque station et grandes gare sur paris

Power BI

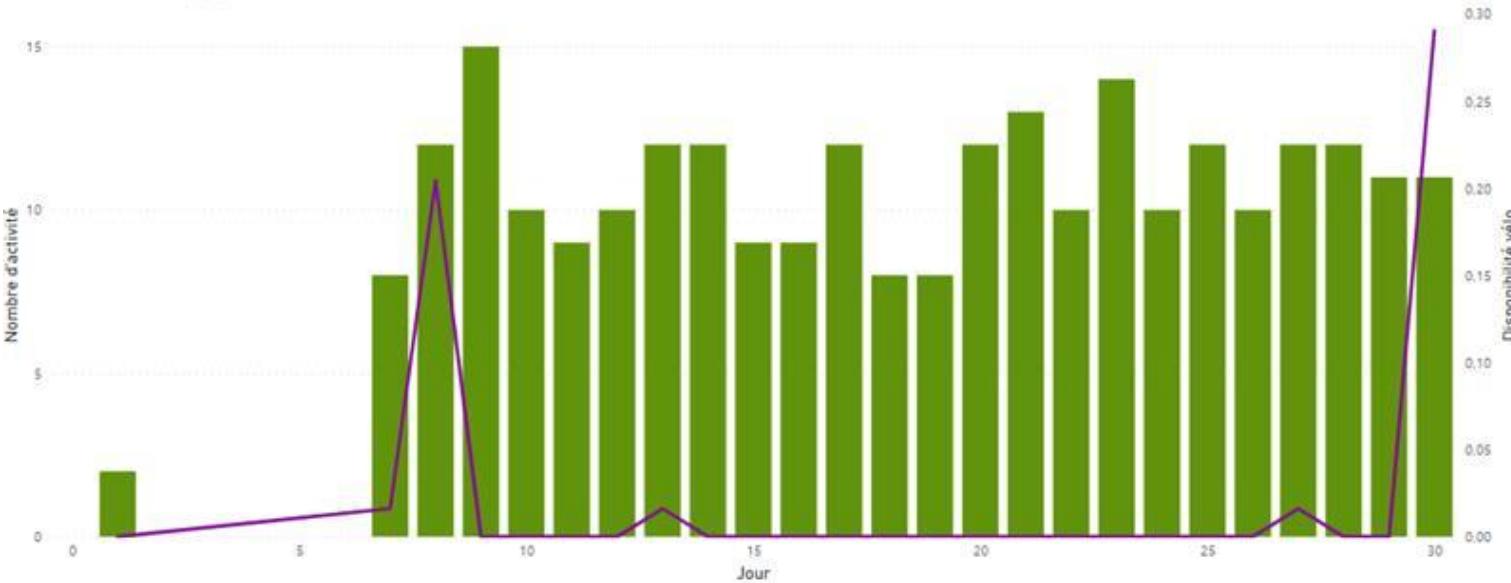
Plusieurs analyses autour de l'activité :

- Disponibilité vélo en fonction des stations
- Activité en fonction des stations
- Capacité des stations
- Disponibilité et activité des vélos en fonctions des stations à proximités des gares

Manque de vélib pour subvenir au besoin de cette station

Disponibilité et Activité des velos par jour pour la station : Place du Maquis du Vercors (Paris)

● Nombre de activité ● dispo_velo

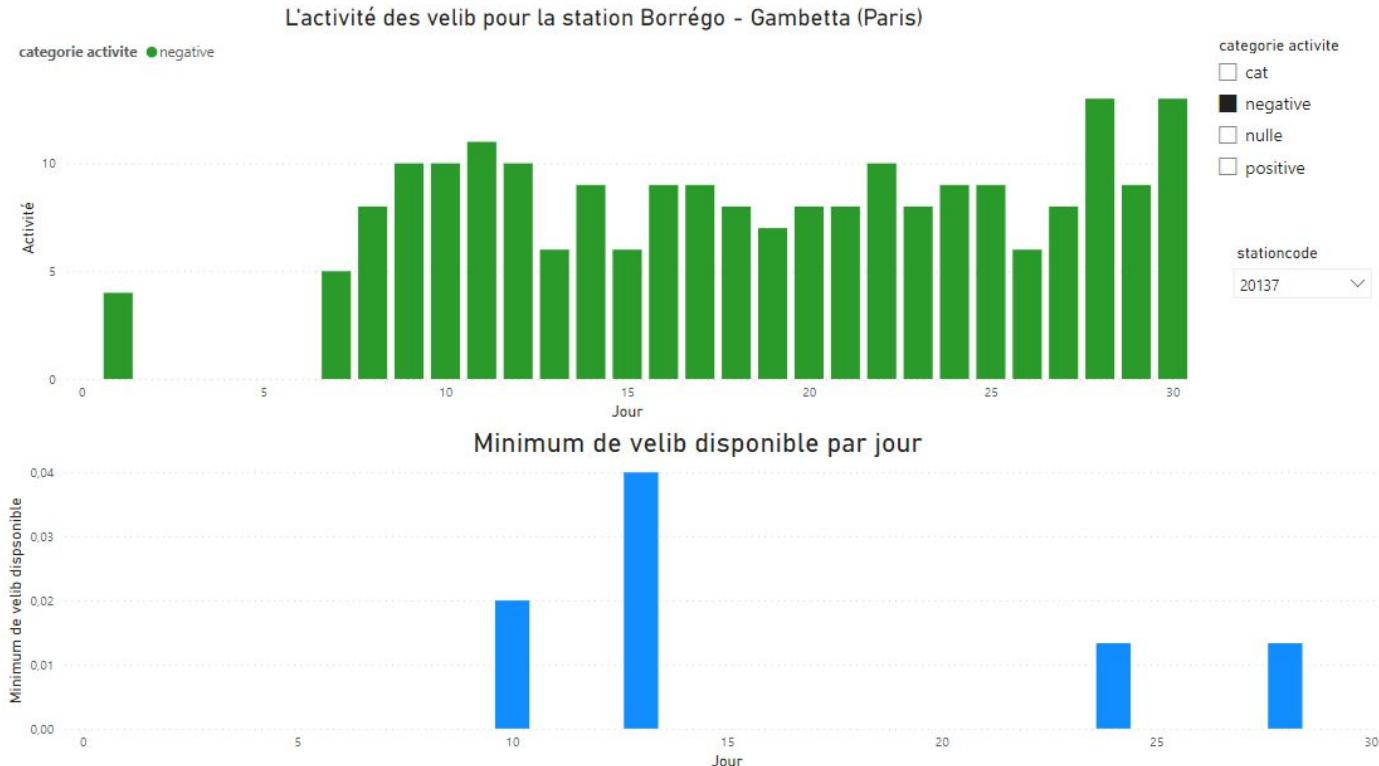


Comptage du nombre de disponibilités de vélos nulles



stationcode	dispo_velo	
0	20019	50
1	20137	41
2	20027	41
3	8049	40
4	20103	39
...
912	23007	1
913	23009	1
914	23010	1
915	19022	1
916	5115	32

Manque de vélib pour subvenir au besoin de cette station



Comptage du nombre de disponibilités de vélos nulles



Manque de vélib pour subvenir au besoin de cette station

L'activité des velib pour la station Le Vau - Maurice Bertaux (Paris)



Comptage du nombre de disponibilités de vélos nulles

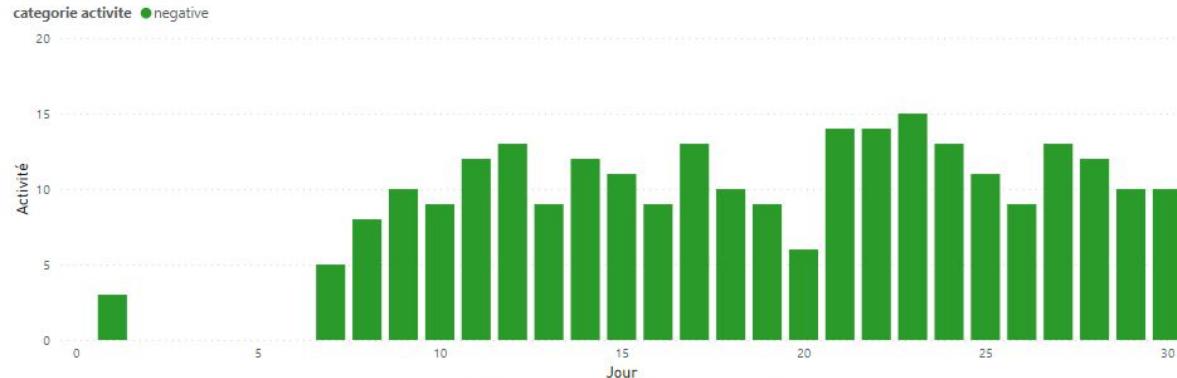


stationcode dispo_velo

0	20019	50
1	20137	41
2	20027	41
3	8049	40
4	20103	39
...
912	23007	1
913	23009	1
914	23010	1
915	19022	1
916	5115	1

Comparaison avec les autres stations sans disponibilités vélo ⇒ Besoin satisfait

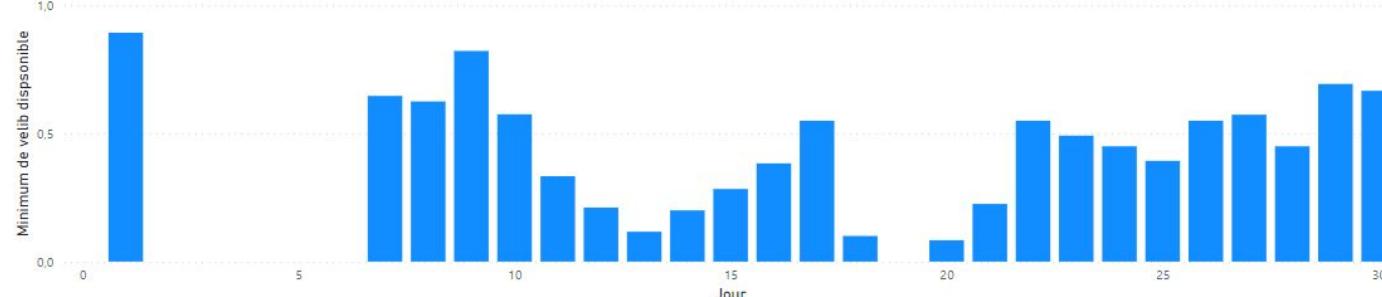
L'activité des velib pour la station Buffon - Jardin des Plantes (Paris)



cat
negative
nulle
positive

stationcode
5115

Minimum de velib disponible par jour

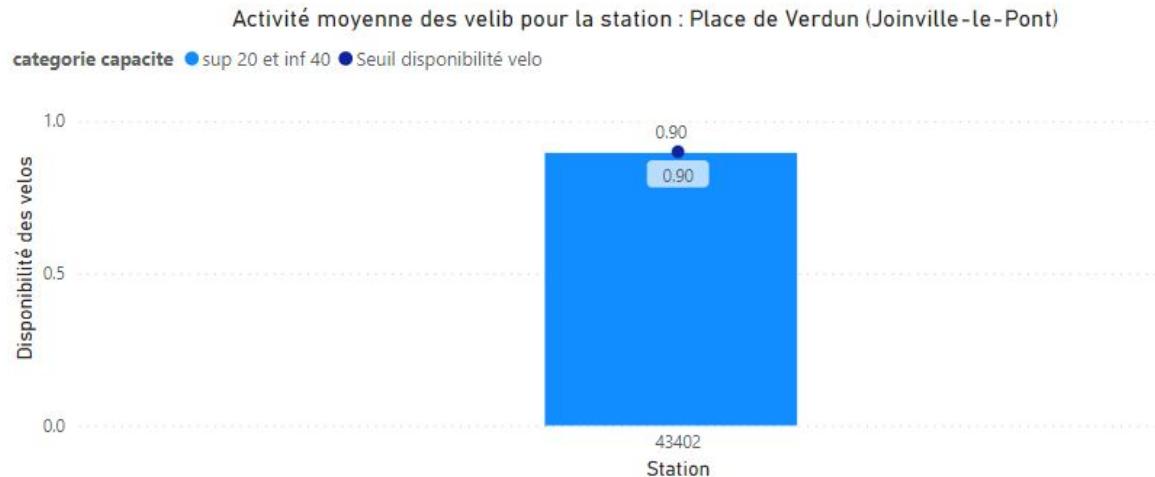


Comptage du nombre de disponibilités de vélos nulles

stationcode ↓

stationcode	dispo_velo	
0	20019	50
1	20137	41
2	20027	41
3	8049	40
4	20103	39
...
912	23007	1
913	23009	1
914	23010	1
915	19022	1
916	5115	1

Dépassement du seuil de vélo disponible à activité égale

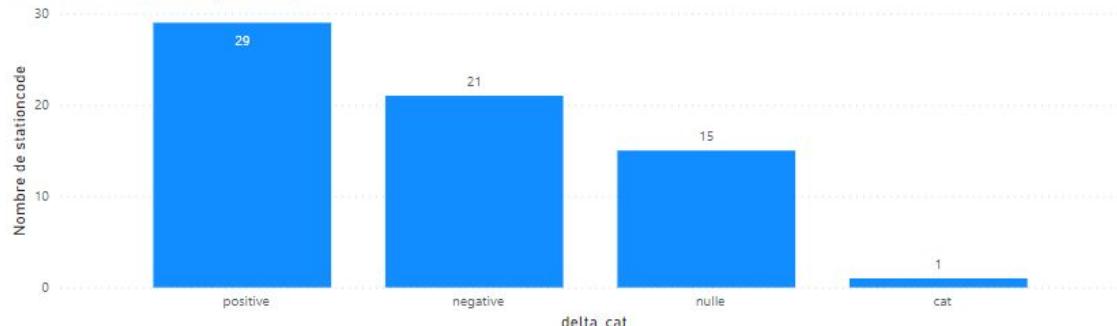


stationcode
11104
11111
11201
12008
12015
12016
12017
12022
12032

date
28 April 2022
29 April 2022
30 April 2022

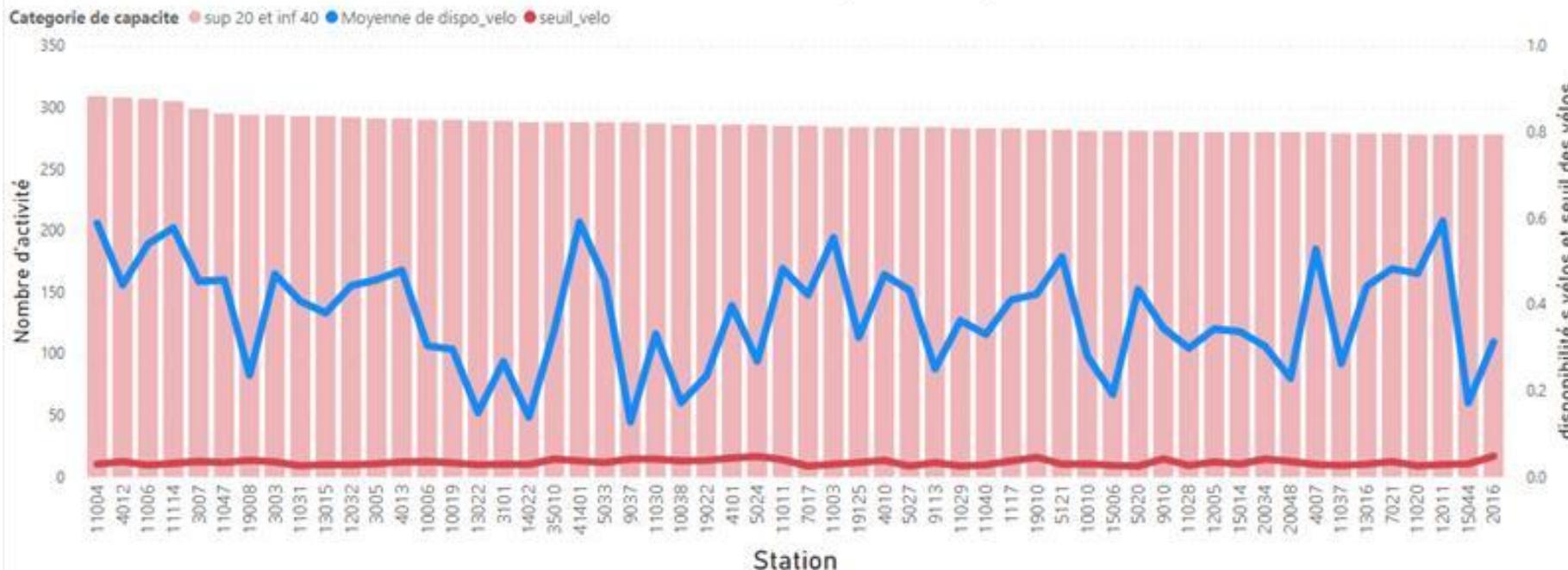
hour
0
1
2
3
4
5
6
7
8

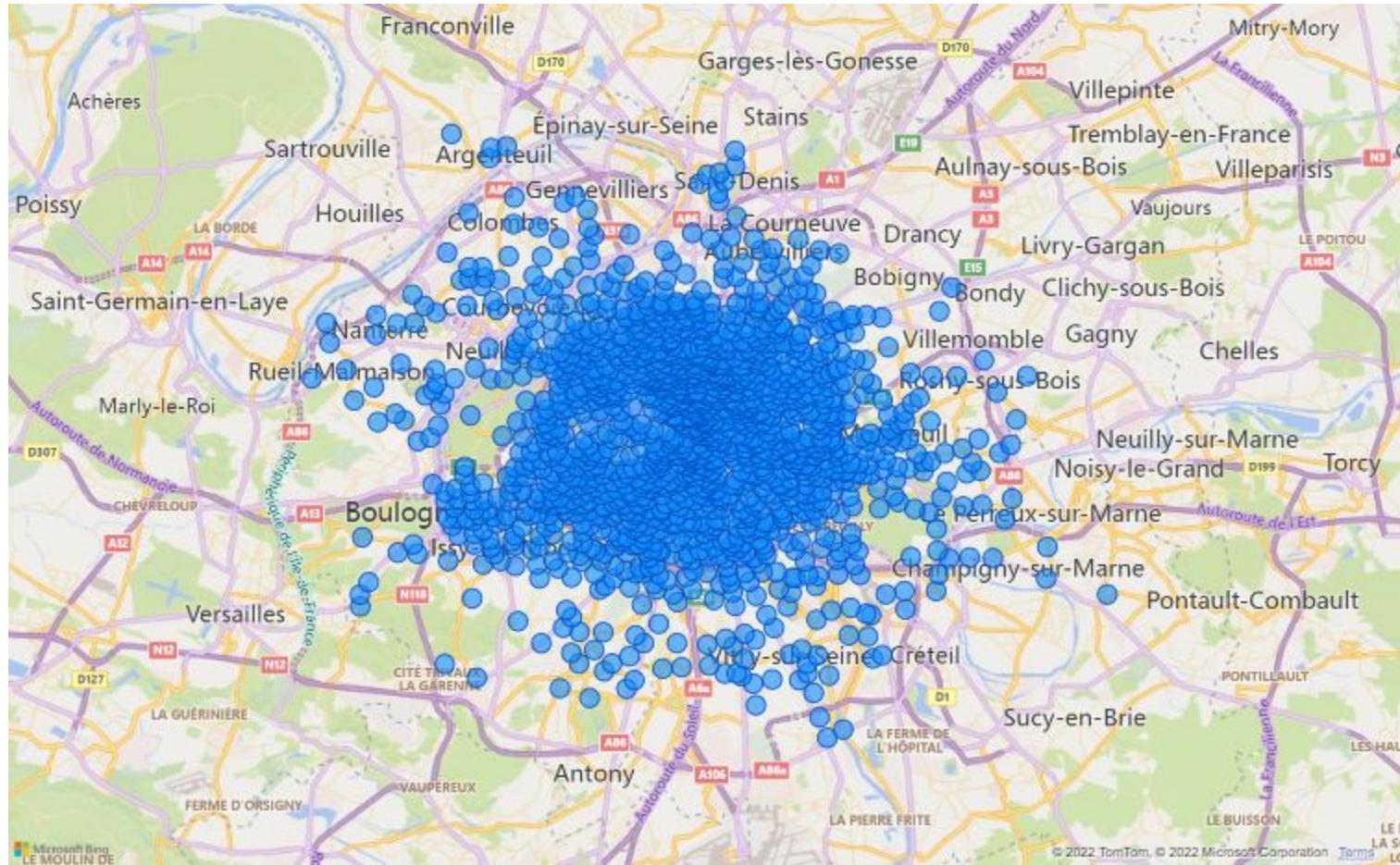
Nombre de stationcode par delta_cat



Activité nulle la plus faible pour la station 11201 à Paris. Dû aux manques de données alors que c'est une station installé

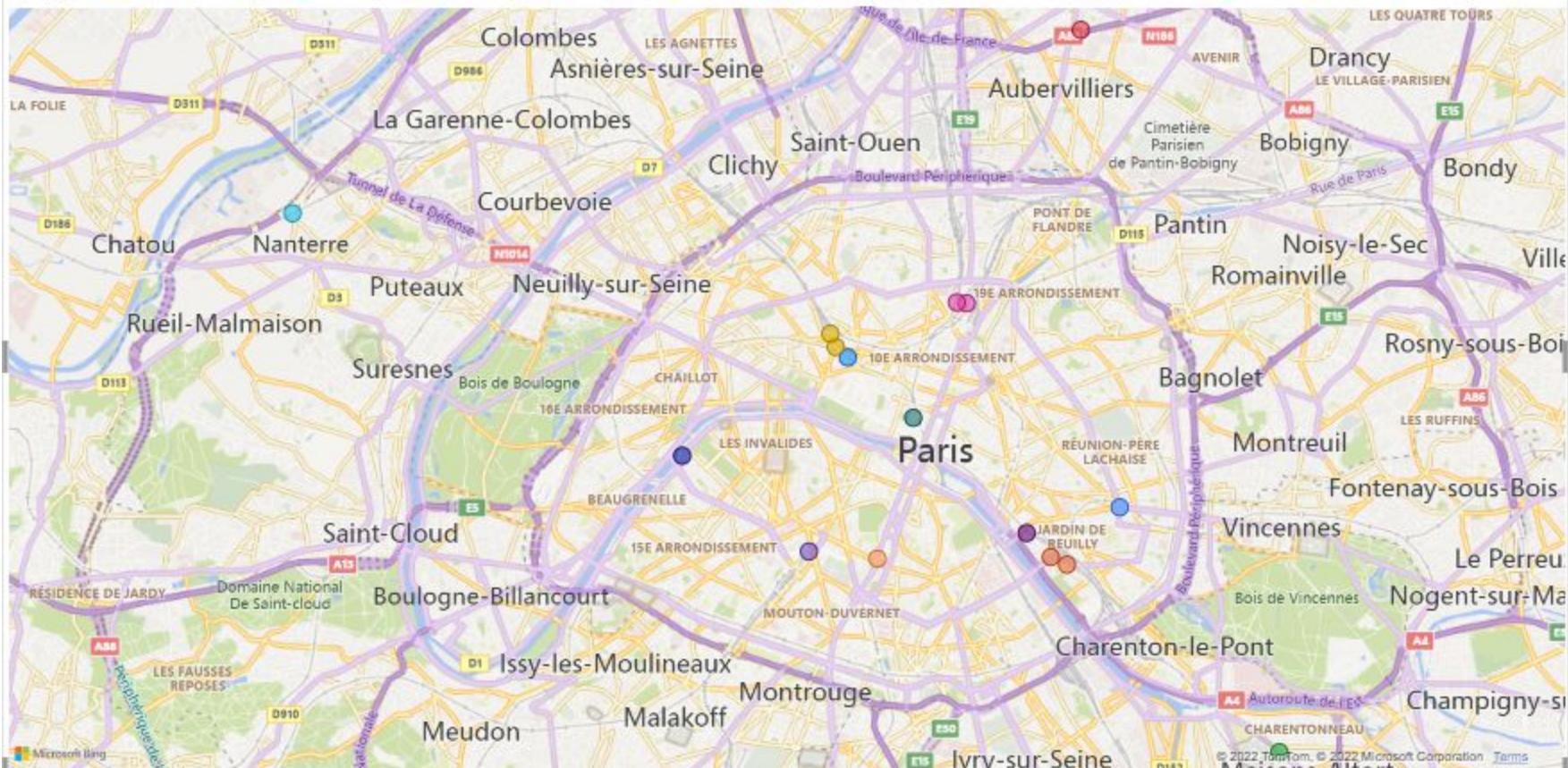
L'activité des stations en fonction des capacité et disponibilité des vélos



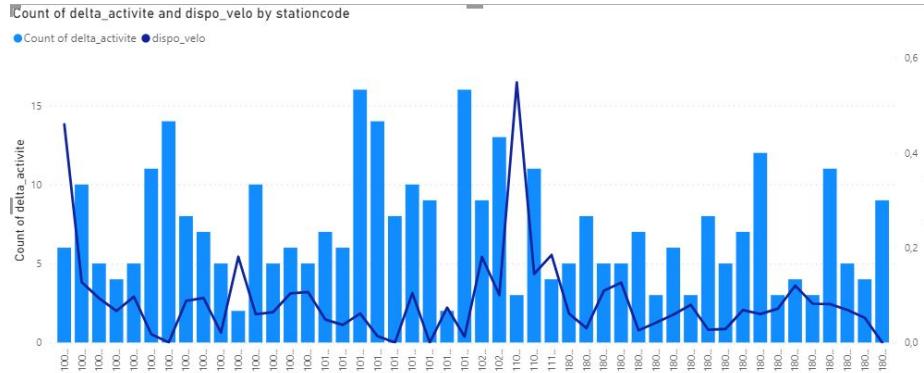


Count of geo_gare by nom_gare, latitude and longitude

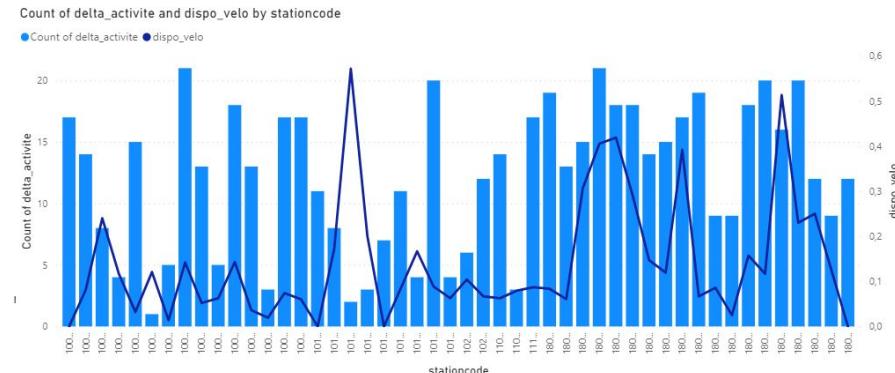
nom_gare ● Auber ● Champ d... ● Gare de B... ● Gare de Lyon ● Gare du ... ● Gare M... ● Gare Sai... ● La Cour... ● Les Halles ● Maisons-... ● Nanterr... ● Nation ● Port Royal



Gare : Gare du Nord
Catégorie activité : activité négative
Catégorie de capacité : supérieur à 20 et inférieur à 40



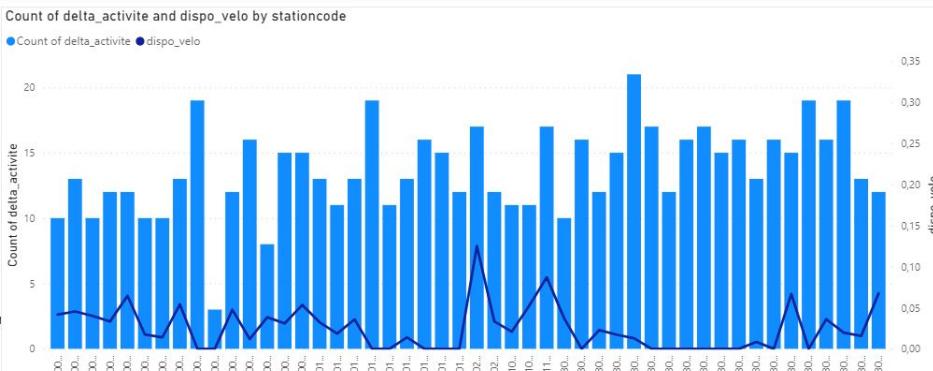
20h



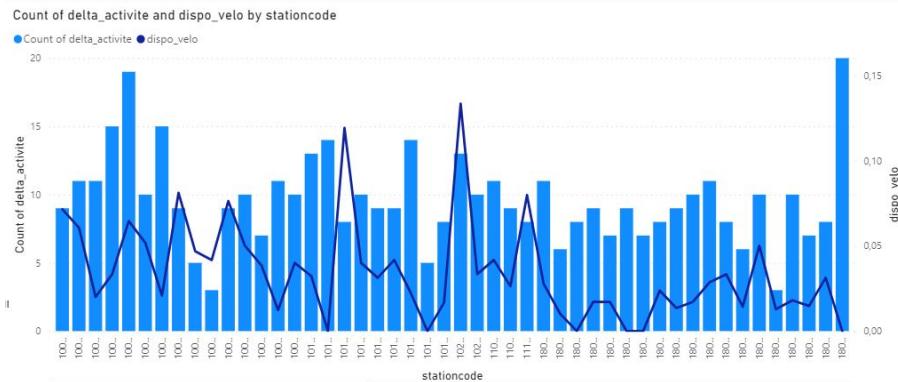
6h

On remarque qu'il y a une mauvaise répartition de vélo sur différent station pour un créneau horaire données

On remarque qu'il y a une mauvaise répartition de vélo sur différent station pour un créneau horaire données



10h



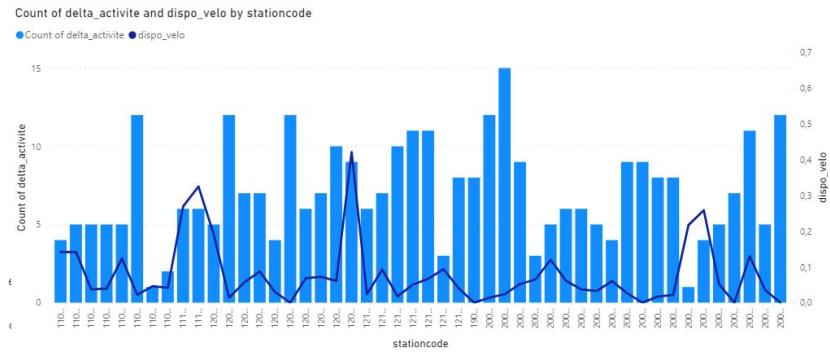
16h

Station : Gare du Nord

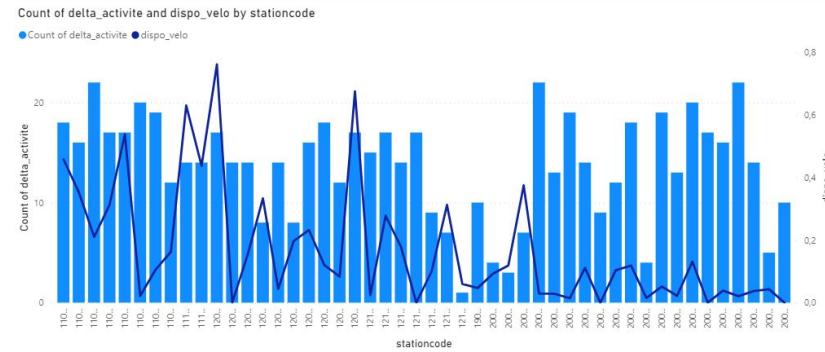
Catégorie activité : activité négative

Catégorie de capacité : supérieur à 20 et inférieur à 40

Gare : Gare du Nord
Catégorie activité : activité négative
Catégorie de capacité : supérieur à 20 et inférieur à 40



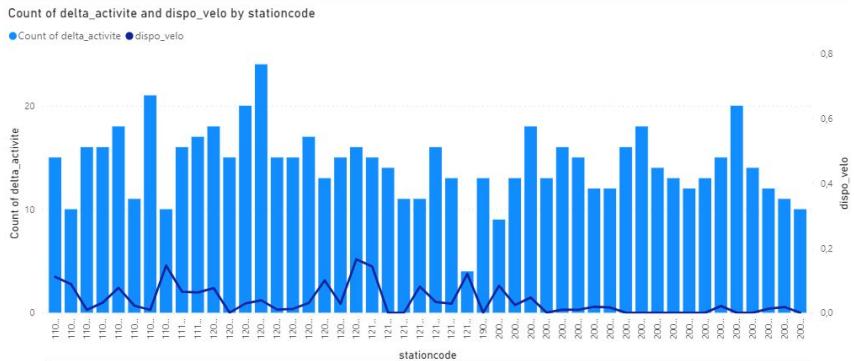
20h



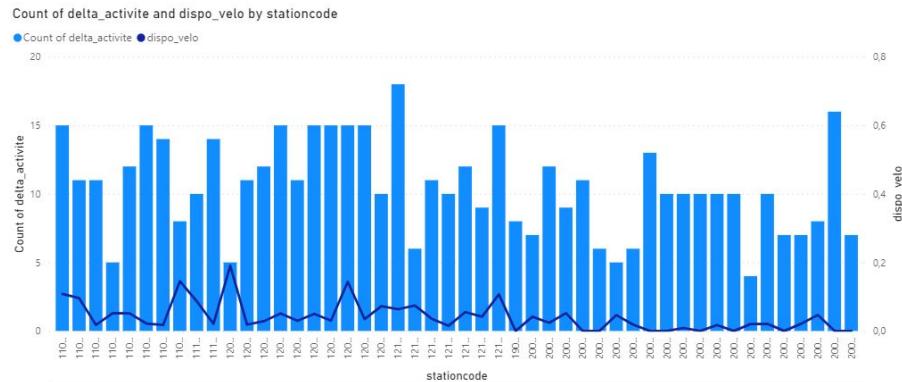
6h

On remarque qu'il y a une mauvaise répartition de vélo sur différent station pour un créneau horaire données

On remarque qu'il y a une mauvaise répartition de vélo sur différent station pour un créneau horaire données



10h



16h

Station : Gare du Nord

Catégorie activité : activité négative

Catégorie de capacité : supérieur à 20 et inférieur à 40