

rapport__

February 23, 2021

1 Qualité de l'air en bretagne

1.1 Contexte du projet

Données

téléchargement des données : [OpenData AirBreizh - Qualité de l'air](#)

Environnement Ce projet a été réalisé en python (sous forme de notebook) à l'aide des bibliothèques : - pandas - numpy - matplotlib.pyplot - seaborn - pyproj

Objectifs

- Analyse des données et interprétation des résultats, en mettant en relation les relevés atmosphériques avec les périodes de confinement.
- Mise en forme des données géographiques en vue d'une cartographie.

1.2 Préparation des données

Dataframes Après retraitement des fichiers téléchargés sur l'API de AirBreizh, nous obtenons deux dataframes qui vont nous permettre de traiter 1. l'analyse de données 2. l'information géographique

	date_ech	lib_zone	confnt	valeur	val_no2	val_o3	val_pm10
2539	2020-01-02 10:00:00	CA Lorient	1	3	1	3	3
2540	2020-01-01 10:00:00	Rennes Métropole	1	3	2	2	3
2541	2020-01-01 10:00:00	CA Golfe du Morbihan - Vannes	1	3	2	3	2
2542	2020-01-01 10:00:00	CA Saint-Brieuc Armor	1	3	2	2	3
2543	2020-01-01 10:00:00	CA Quimper Bretagne Occidentale	1	3	2	3	2
2544	2020-01-01 10:00:00	CA Lorient	1	3	2	3	2
2545	2020-01-01 10:00:00	CA du Pays de Saint-Malo (Saint-Malo Aggloméra...	1	3	1	2	3
2546	2020-01-01 10:00:00	Brest Métropole	1	3	2	3	2

	code_zone	lib_zone	geom	Longitude	Latitude
zone					
4	200069409	CA Saint-Brieuc Armor	POINT (269153.2682993477 6832824.934141148)	-2.832059	48.452434
1	200067932	CA Golfe du Morbihan - Vannes	POINT (269961.1106939723 6746307.8992528645)	-2.736238	47.676516
3	200068120	CA Quimper Bretagne Occidentale	POINT (172753.63056731108 6795400.543816281)	-4.085219	48.045952
6	243500139	Rennes Métropole	POINT (349555.0070453745 6789093.386528317)	-1.711339	48.108092
5	243500782	CA du Pays de Saint-Malo (Saint-Malo Aggloméra...	POINT (338719.75810268405 6844013.788199687)	-1.902423	48.595360
0	242900314	Brest Métropole	POINT (147878.22420658765 6837175.794037035)	-4.470520	48.399641
2	200042174	CA Lorient	POINT (229772.04043578357 6768639.926683848)	-3.293846	47.849403

Le premier contient 7 colonnes traitant les informations des 2547 mesures.

La colonne `date_ech` donne la date à laquelle a été prise la mesure. Les dates s'étendent du 01-01-2020 au 20-01-2021.

La colonne `lib_zone` contient le libellé de la zone pour laquelle est enregistrée la mesure. Elle comporte 7 valeurs différentes (pour les 7 agglomérations principales de la région) Cette colonne est commune aux deux Dataframes et permettra éventuellement des jointures pour une utilisation ultérieure.

Une colonne `confinnement` a été ajoutée pour noter les périodes de confinement (`confnr` est à 1 si la date correspond à un jour de confinement, 0 sinon).

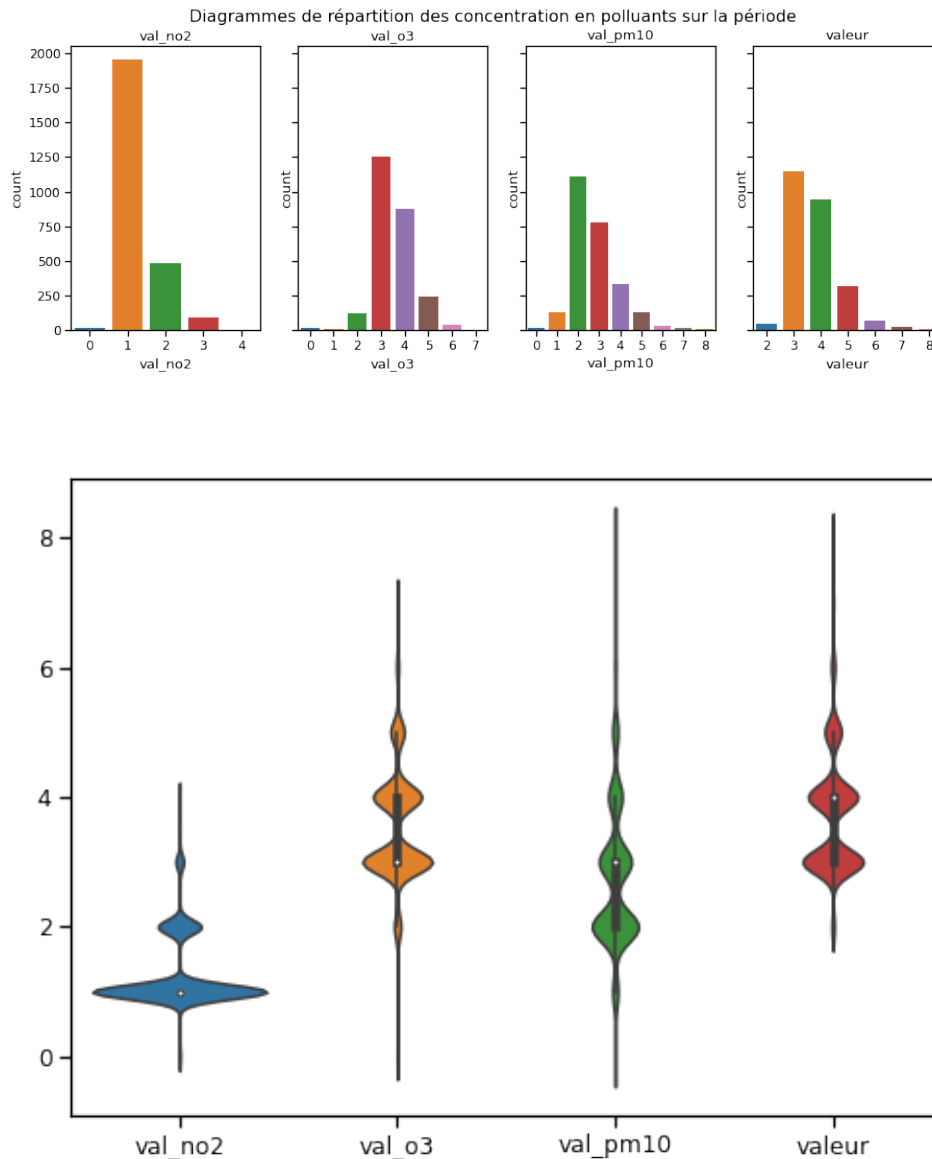
La `valeur` correspond à l'indice de qualité de l'air :

	qualif
valeur	
2	Très bon
3	Bon
4	Bon
5	Moyen
6	Médiocre
7	Médiocre
8	Mauvais

Les 3 autres colonnes `val_no2`, `val_o3` et `val_pm10` correspondent respectivement aux l'indices de pollution au dioxyde d'azote, à l'ozone, et aux particules fines.

1.3 Analyse des données

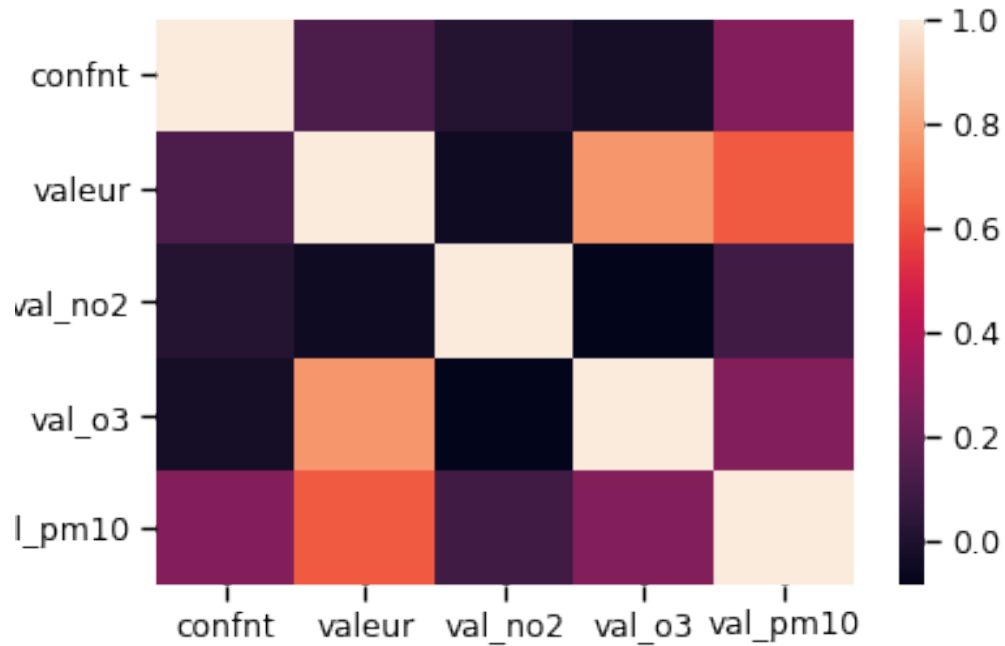
Nous débutons cette étude par une visualisation simple des distributions. Puis on s'intéressera aux corrélations pouvant exister entre les variables, et nous visualiserons les données de manière exhaustive.



Pour les concentrations en polluants comme pour la valeur de l'indice de qualité de l'air, la distribution des valeurs est unimodale, avec une asymétrie à gauche. Il est également intéressant de noter que la concentration en dioxyde d'azote (NO₂) varie peu et présente presque 70% du temps la valeur 1.

Il est également intéressant de regarder si les variations des concentration en polluants ont un lien entre elles et avec celle l'indice de qualité de l'air.

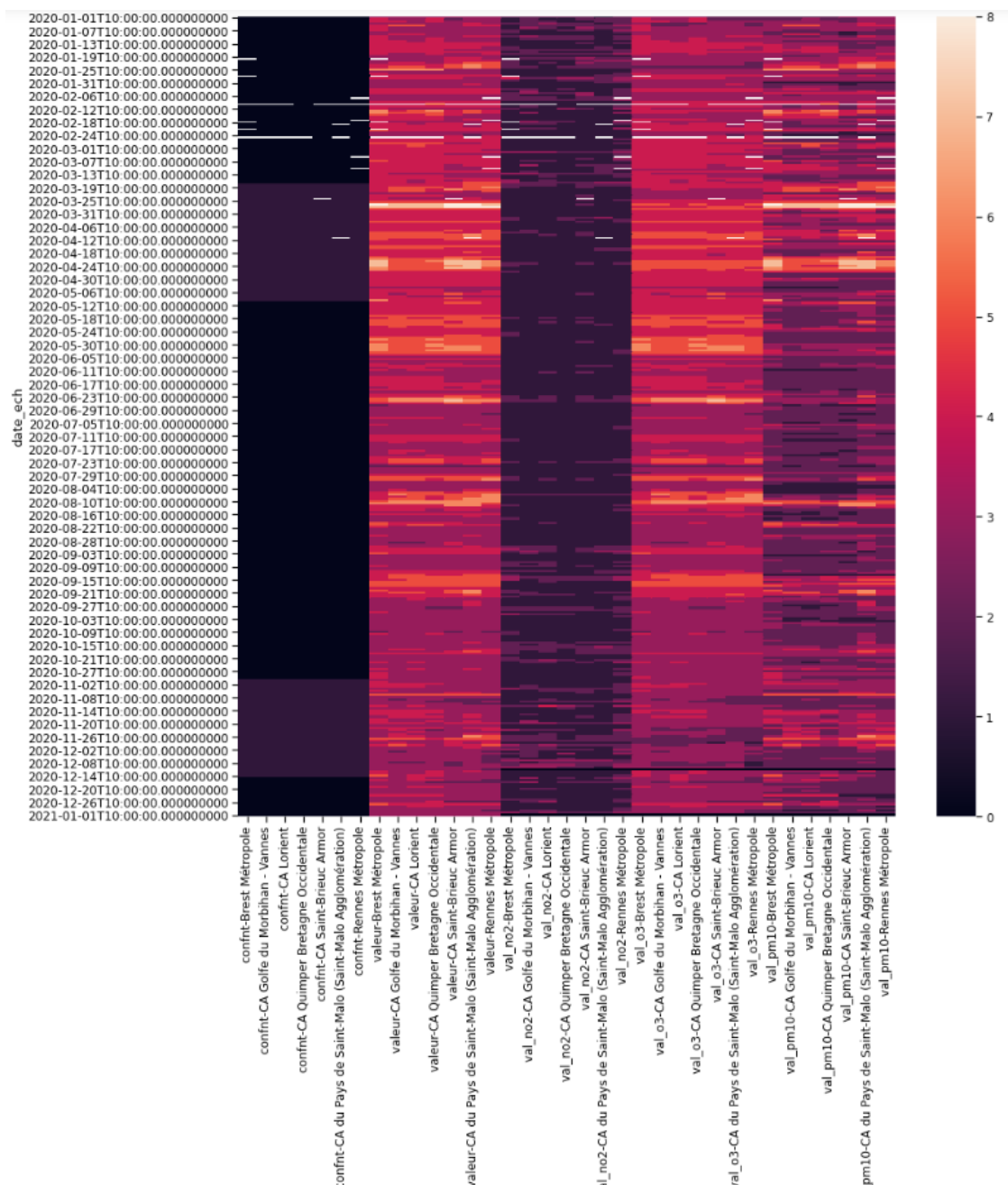
Comme nos variables sont discrètes, la visualisation des nuages de points ne donnera rien d'intéressant. Par contre, on peut utiliser une matrice de corrélation, ou "Heatmap" pour mettre en avant les liens éventuels.



Les corrélations qui apparaissent nettement sont celles qui existent entre la valeur de l'indice de la qualité de l'air (bon / moyen / médiocre...) et les concentrations en ozone (O3) et en dioxyde d'azote (NO2). La corrélation de ces deux valeurs *entre elles* n'est par contre pas nette.

Par ailleurs, cette carte indique qu'il n'y aurait pas, a première vue, de lien de cause à effet entre confinement et indice de la qualité de l'air.

On peut voir tout cela plus en détails dans la matrice de corrélation ci-dessous.



Cette carte très complète nous montre les valeurs de l'indice de la qualité de l'air et des concentrations en polluants pour les différentes zones et à travers le temps. Afin d'en faciliter la lecture, on a également choisi d'y faire apparaître les périodes de confinement.

Il en ressort que : - les corrélations définies à l'aide de la première matrice sont confirmées. La concentration en Ozone (O3) particulièrement évolue de manière très similaire à l'indice de la qualité de l'air. Par ailleurs, on assiste à une importante dégradation la qualification de l'indice lors des pics de pollution aux particules fines (voir dans la matrice ci-dessus les pics de fin mars et mi-avril 2020). On peut en conclure que **les pollutions à l'ozone et aux particules fines sont déterminantes dans la définition de l'indice de la qualité de l'air.**

- On observe que les deux pics de pollution en PM10 et O3 les plus importants de

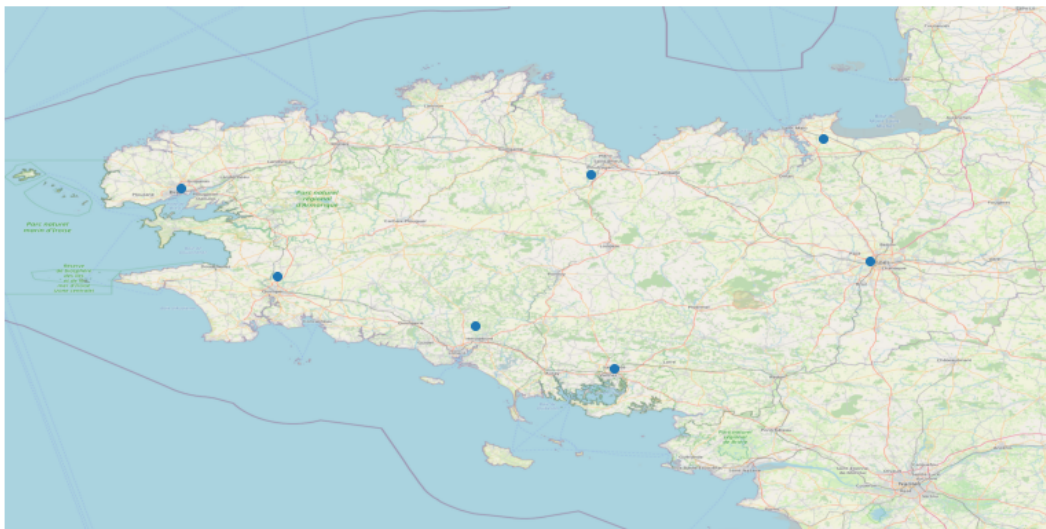
2020 (donc dégradation de l'indice, comme vu ci-dessus) se situent tous les deux pendant le premier confinement. On peut supposer que les pollutions domestiques ont remplacé celles du trafic routier en termes d'émission de particules fines et de production d'ozone. Ces données seraient également à croiser avec des données météorologiques, les précipitations favorisant par exemple la dispersion des particules fines, et l'ensoleillement la production d'ozone, etc.

- La concentration en dioxyde d'azote (NO₂) varie effectivement peu, et son augmentation ne semble pas liée au confinement, mais peut-être **plutôt à la saison** : basse en été, augmentations épisodiques en hiver.
- Enfin, la répartition entre départements est globalement homogène. On observe des concentrations en polluants légèrement plus élevées sur les agglomérations de Brest et Saint-Malo.

1.4 Préparation des données cartographiques

Afin de traiter les données cartographiques, un retraitement des coordonnées a été nécessaire pour les passer du système de coordonnées Lambert93 en coordonnées GPS standards (wsg84).

Ce retraitement a permis de créer la carte ci-dessous, avec des points indentifiés qui peuvent permettre la visualisation de données.



L.M.