

Rapport Data Visualisation

AZOUÏ Aymen BOUICHE Ryma SAIDJ Lydia

TOINON Vincent

January 2021

1 RÉSUMÉ

Dans le cadre de l'UE Data Visualisation, nous avons été amenés à réaliser un projet de visualisation développé en d3.js. Nous avons décidé de faire une visualisation montrant une comparaison de taux de pollution causée par l'avion selon de différents polluants dans les régions de France. Nous nous sommes intéressés à ce thème car l'aviation est l'une des activités humaines qui génèrent un taux important de pollution. L'aviation mérite une attention particulière.

2 INTRODUCTION

Le transport aérien, comme toute autre activité, génère une pollution atmosphérique aux échelles locale, régionale et planétaire. Toutes les mesures réalisées autour des aéroports montrent que les concentrations en polluants sont supérieures aux concentrations mesurées au cœur des grandes villes.

3 DONNÉES

La source de donnée principale de ce projet est le site internet [1], on s'est intéressé sur les années entre 2018 et 2020. Notre dataset contient 40836 observations, et 5 colonnes (date,region,pm10,pm 25,o3).

4 INFRASTRUCTURES TECHNIQUES

Les technologies utilisées dans ce projet :

- Jupyter pour le nettoyage de données -i suppression de colonnes inutiles, ajout de colonne, renommage de région et concaténation de tables.
- html /css /javascript / bootstrap : pour la création de l'interface web
- D3.js pour la mise en place des visualisations.

5 TYPES DE VISUALISATION

L'objectif de ce projet étant de pouvoir extraire de l'information de manière visuelle, il est très important d'imaginer les différentes visualisations possibles pouvant permettre cela.

5.1 VISUALISATION 1

* Visualisations orientées cartes géographiques:

Intuitivement et dès que l'on s'intéresse à des données relatives à une zone géographique l'une des premières idées à exploiter est de projeter les données sur une carte (la France en l'occurrence ici comme le montre la figure 1).

Les avantages de ce type de visualisation:

- Comparaison des régions entre elles et avec le référentiel national
- Possibilité de visionner les données de façon globale et synthétique.

La conjugaison du paramètre "couleur" met en évidence un changement au niveau de taux des gaz émis. La légende permet de bien comprendre la visualisation.

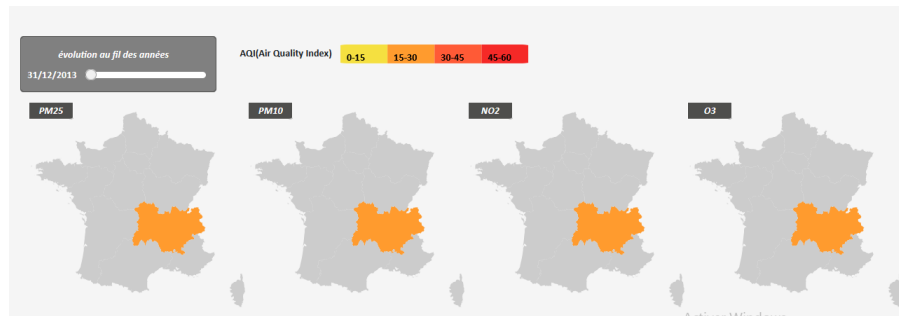


Figure1

5.2 VISUALISATION 2

* Visualisation orientées histogrammes, diagrammes ou lignes

Contrairement aux cartes qui permettent de comparer les régions entre elles et expriment de façon limitée la dimension temporelle, les diagrammes permettent de mettre en exergue l'évolution dans le temps de façon plus expressive. Etant donné que nos données vont de 2014 à 2020 nous pouvons analyser les taux des gaz pour chaque région.

Nous montrons un exemple sur la visualisation 2 dans l'image ci-dessous en choisissant la région **alpes_cote_d'azur_est**.

Presentation visu

alpes-cote-dazur-est

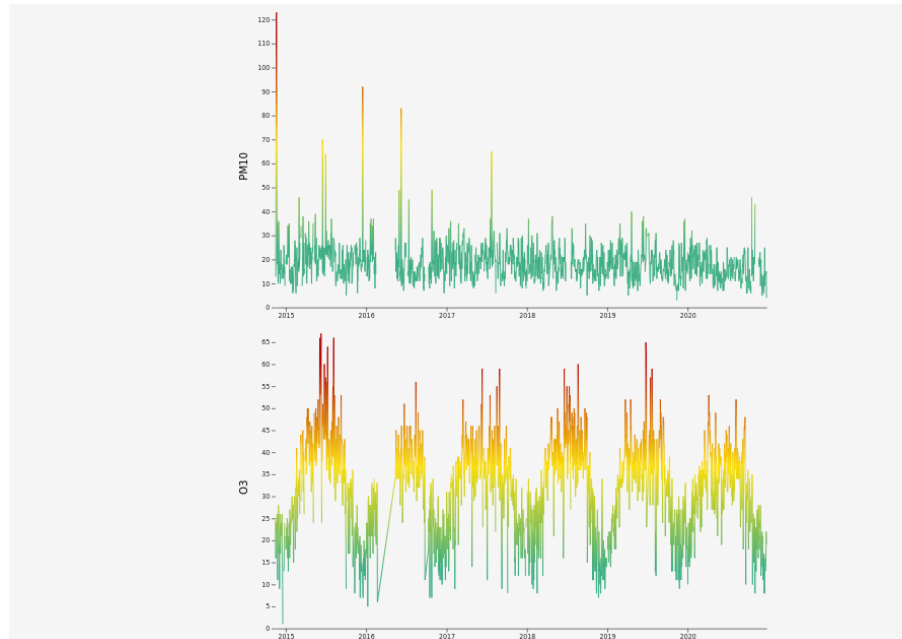
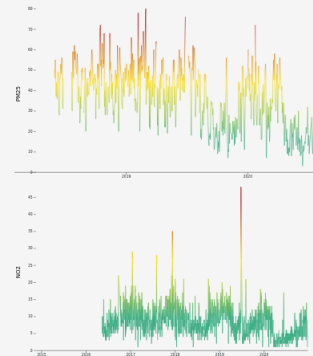


Figure2

Nous avons ainsi la possibilité d'agrandir le schéma afin de mieux visualiser l'évolution du taux du polluant sur un laps de temps restreint.

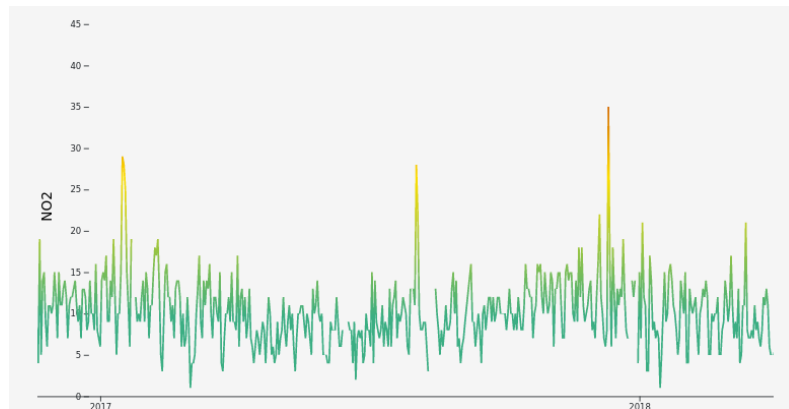


Figure3

5.3 VISUALISATION 3

Cette visualisation sert à calculer les émissions de gazs à effets de serre (GES) produites par un vol regional à l'interieur du territoire français en fonction d'une ville de départ, une ville d'arrivée et d'un modèle d'avion.

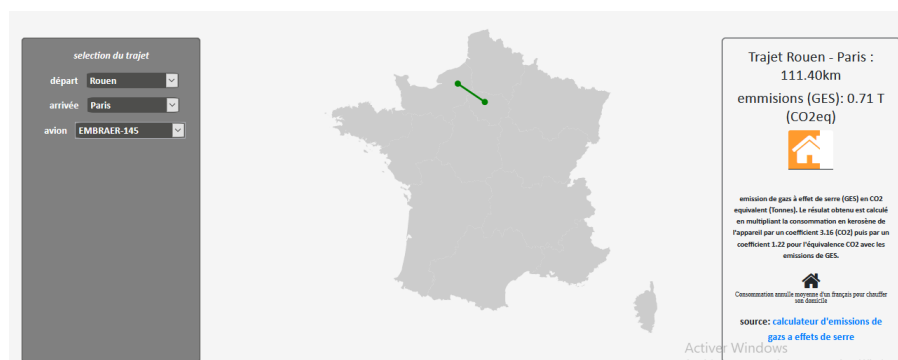


Figure4

6 À QUI S'ADRESSE CETTE VISUALISATION?

L'utilisateur visé par cette étude peut être:

- Tout utilisateur intéressé par la pollution atmosphérique causée par l'aviation.
- Une personne travaillant au sein d'une compagnie aérienne, par exemple chargée de l'analyse de taux de pollution dans l'air.

Dans le cas du deuxième utilisateur, ces visualisations et les interprétations peuvent servir de base ou de travail préliminaire pou-

vant conduire à des études complémentaires impliquant d'autres jeux de données (type de moteurs d'avion, type d'avion utilisant les nouvelles techniques à l'atterrissage et au décollage diminuant la pollution d'aire ... etc.).

La disposition des visualisations au sein de l'interface web du projet permet de développer un raisonnement qui satisfait à la fois les utilisateurs avec un profil technique ou non. Le besoin peut être différent mais la logique suivie pour organiser les visualisations permet d'utiliser avec une meilleure compréhension du taux de polluants causé par l'aviation en France à plusieurs niveaux:

1. Géographique
2. Temporal

7 À QUELLE QUESTION LA VISUALISATION PERMET-ELLE DE RÉPONDRE?

Cette visualisation permet de répondre à la question suivante : Quelles sont les régions ayant le taux le plus/moins élevé de pollution dû à l'aviation?

8 OPTIMISATION DE LA VITESSE DE COMPRÉHENSION DES VISUALISATIONS

Afin d'accélérer le processus de compréhension des visualisations nous avons joué sur plusieurs paramètres d'IHM (interface homme-machine) nous avons centré les visualisations en les faisant défiler verticalement , ce qui correspond au raisonnement que nous voulons voir suivre l'utilisateur.

L'utilisation de composants graphiques fournissant un retour visuel immédiat tels que des barres de défilement, des boutons de réglage de l'affichage accélèrent l'affichage des informations recherchées par l'utilisateur.

Les Mouseovers et les clics permettent à l'utilisateur de survoler les informations plus rapidement ou de les définir pour une comparaison ou une analyse plus approfondie

9 LIMITES

En ce qui concerne les limites:

- Pour la première visualisation, l'idée du départ était de faire une représentation pour toute la carte du monde, mais nous n'avons malheureusement pas trouvé un jeu de données qui couvre tous les pays du monde, nous nous sommes donc focalisé sur la carte de France.

- Nous avons voulu faire une visualisation qui représente le taux d'évolution pour chaque polluant ainsi que leur moyenne, en fonction de la l'année pour chaque région, dans le but de comparer les régions de France selon leur pollution.

10 MÉTHODE DE TRAVAIL

Pour la réalisation de ce projet nous avons utilisé :

- Discord pour la communication, transfert des ressources et documents ainsi que le partage d'écran qui nous a été très utile pour le travail en Peer Programing.
- Google doc et overleaf pour la réduction de peer review et le rapport.

11 CONCLUSION

Cette étude nous a permis de tout mettre en œuvre pour extraire un maximum d'information des données de façon visuelle. Le challenge est donc de changer les points de vue et de penser différentes visualisations afin d'apporter le plus d'éclairage possible. La mise en place de cette description théorique du projet met en évidence l'importance de la phase nettoyage/traitement de données et du choix de l'indicateur de qualité sans lesquels les visualisations perdraient de leur valeur. Pour bien réussir ce projet, nous nous inspiré des exemples se trouvent sur le site Observable[2].

References

- [1] <https://aqicn.org/data-platform/register/?fbclid=IwAR3zROhce7YKS473CZb0mpBUr8lehZ>
- [2] <https://observablehq.com/@d3>.