

Concours Assises du Transport Aérien - mai 2018

Gwilherm Baudic

Introduction

Ce rapport vise à présenter de façon succincte ma proposition de datavisualisation pour les Assises du Transport Aérien.

Démarche

Au vu des critères énoncés dans le règlement du concours, le choix d'une interface de visualisation web est apparu comme une évidence. J'ai choisi de m'appuyer sur des cadres et bibliothèques *open source* reconnus et que j'avais déjà pour la plupart manipulés par le passé : Angular 5, Bootstrap pour le style général, Leaflet pour l'affichage cartographique et Chart.js pour l'affichage de graphiques. Le caractère libre ainsi que le support existant pour ces différents composants est un atout pour pérenniser le code proposé ici. De plus, les composants retenus sont tous prévus pour une utilisation *responsive*, nécessaire pour permettre une adaptation à toutes les plateformes.

Principe

Le visiteur arrive sur une carte de France métropolitaine sur laquelle sont placés les aéroports français. Cette visualisation géographique du maillage du territoire permet de répondre au défi n°3.

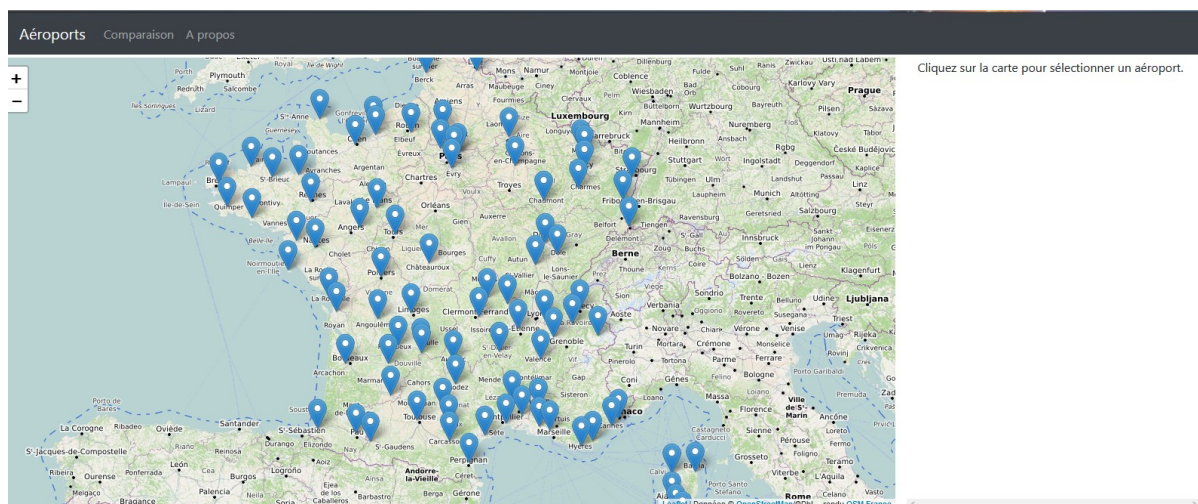


Illustration 1: Vue générale

Il suffit alors de cliquer sur le marqueur d'un aéroport donné pour obtenir des informations complémentaires : nombre de passagers et de destinations pour la dernière année considérée, top 10 des destinations par nombre de passagers, répartition des

passagers par faisceau (court-moyen-long-courrier), pollutions constatées pour 2016 (les fichiers des années précédentes ne concernant en effet que les aéroports parisiens), évolution du nombre de passagers sur la période des données. On constate ainsi dans l'exemple ci-dessous (Toulouse-Blagnac) que les principales destinations sont, sans grande surprise, les aéroports parisiens de Roissy et Orly. Pour les aéroports où les données ne sont pas disponibles, cette absence de données est clairement indiquée.

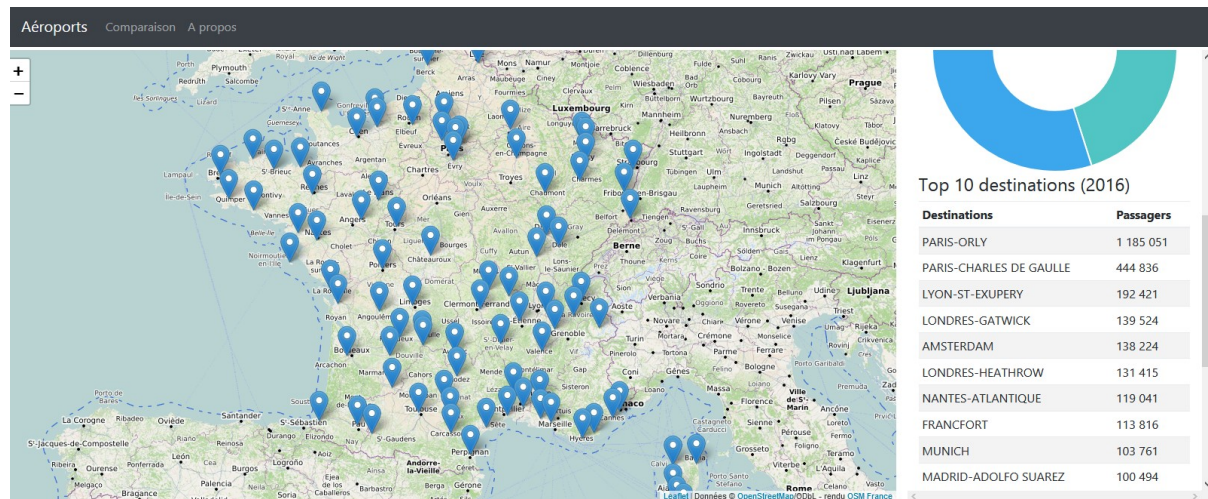


Illustration 2: Top 10 des destinations pour Toulouse-Blagnac (LFBO)

En cliquant sur le bouton “Plus d’infos” situé en bas, on arrive à un historique par type de vol du nombre de passagers sur la période couverte par les données (1990-2016), mis en parallèle avec le nombre de vols, pour chaque type de faisceau considéré. S’il le souhaite, l’internaute peut choisir une période temporelle plus restreinte, avec une précision à l’année. Cette vue permet de répondre aux défis 1, 2 et 4.

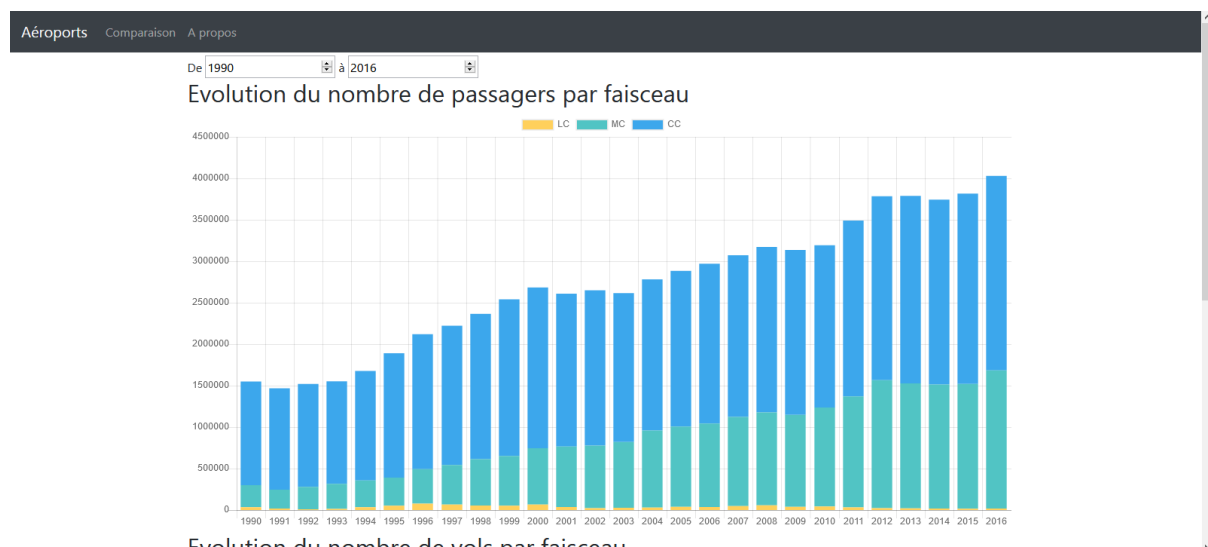


Illustration 3: Evolution du nombre de passagers par faisceau pour Toulouse-Blagnac

Cependant, cette vue ne couvre qu’un aéroport donné. Il pourrait être intéressant de

voir si une évolution (par exemple, une augmentation de trafic) est spécifique à un aéroport ou si elle est généralisée. A cet effet, une vue comparative est proposée. Elle reprend les mêmes informations que la vue précédente, mais avec cette fois les données pour deux aéroports placées côte-à-côte. Là aussi, il est possible de restreindre la période temporelle affichée.

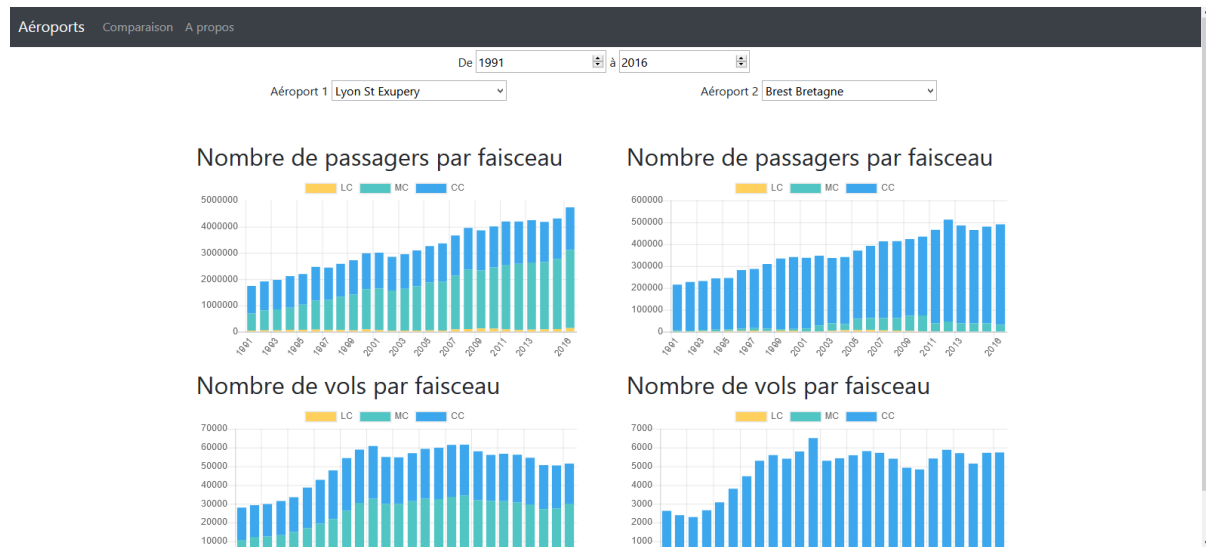


Illustration 4: Comparaison des données entre les aéroports de Lyon et Brest

Enfin, la source ainsi que les limites des données fournies sont indiquées à l'internaute par une page "A propos", visible depuis toutes les vues, traitant ainsi le défi n°5.

Traitements de données effectués

Les fichiers fournis étant proposés au format CSV, nous avons choisi de les traiter au moyen du logiciel de statistique R. Les sorties sont au format JSON pour permettre une intégration directe à l'interface développée. Bien que l'inclusion directe des données au sein du code Javascript généré entraîne une légère augmentation de sa taille, cela nous a permis d'avoir une interface web pouvant s'exécuter seule, sans avoir à développer en plus un mini-serveur pour nous renvoyer les données.

- Fichier aéroport : nous avons extrait uniquement les aéroports de France métropolitaine, avec leur nom, codes OACI et IATA ainsi que leur position latitude/longitude pour un affichage sur le fond de carte.
- Fichiers pollution : nous avons effectué une fusion par faisceau (court/moyen/long-courrier) pour chaque aéroport, uniquement sur le fichier de 2016 car les années précédentes ne couvrent que les aéroports parisiens.
- Fichiers retards : non pris en compte dans ce projet, en raison d'une période temporelle différente des autres fichiers (début en 2012 contre 1990)
- Fichiers trafic : calcul du nombre de passagers et de vols par année, par aéroport, par faisceau. Nous avons également établi un top 10 des destinations par nombre de passagers pour la dernière année des données (2016 dans cette phase 1), et compté le nombre de destinations desservies par aéroport, toujours pour cette dernière année.

Pour tous les traitements, les données sont volontairement sommées par année pour simplifier, bien que les fichiers fournis offrent une précision au mois. Le fond de carte est issu d'OpenStreetMap et est utilisé tel quel avec attribution.

Conclusion

Nous avons réalisé une visualisation permettant, à partir du choix d'un aéroport donné, d'en avoir à la fois une description succincte (activité, pollutions) et de voir les évolutions de l'activité sur la période temporelle couverte par les données fournies. Une vue comparative a également été développée. Nous avons fait le choix d'une vision synthétique (données par an et non par mois, aucun traitement statistique) dans l'objectif de rester accessible au plus grand nombre et de proposer une interface intuitive, ce qui est un des objectifs du concours. Enfin, il convient de noter que la solution proposée traite les cinq défis énoncés.

Il serait possible d'étendre les travaux présentés ici. Tracer les différentes liaisons depuis un aéroport donné en reflétant leur importance par des lignes plus ou moins épaisses est une possibilité, qui permettrait en outre de valoriser les données sur les retards non prises en compte dans notre proposition. L'usage d'une source externe de données pour compléter les descriptions des aéroports, par exemple nombre de pistes et longueurs, pourrait également être intéressant.