



CONCOURS DE DATA-VISUALISATION DES DONNEES DE TRANSPORT AERIEN

Avril/mai 2018

Mr. Hirbec Jordan
Mr. Kermouni Vincent

08/05/2018

Ce document vise à présenter les résultats de la conception d'une interface innovante, interactive et pédagogique visant à visualiser en temps réel les informations relatives aux transports aériens fournies dans le cadre de ce concours. Cette application, se veut élégante et concise. Les graphiques sont interactifs, dynamiques et exploitables par tous les utilisateurs du site du ministère de l'environnement.

La première étape du projet consistait à prendre connaissance des données afin d'effectuer une étape de data pre-processing permettant d'obtenir un jeu de données agrégé et nettoyé. Par exemple, durant cette phase nous avons joint tous les fichiers en quatre afin d'obtenir une vue plus globale et pouvoir effectuer des conditions (filtres) par la suite. De plus, lors des analyses de corrélation les variables qualitatives ont été encodées. Pour les données manquantes nous n'avons pas effectué de remplacement de type (moyenne, médiane ou constante), nous avons simplement garder les données présentes et affiché le nombre de données manquantes pour chaque fichier dans l'onglet supervision.

La solution a été conçue en python, avec l'aide de la librairie de visualisation interactive : Bokeh (Open Source). Bokeh permet d'être lancé au travers d'un navigateur internet pour obtenir des présentations en ligne. L'utilisateur peut ainsi se rendre dans son navigateur favori et visualiser de façon simple et dynamique des informations relatives au transport aérien.

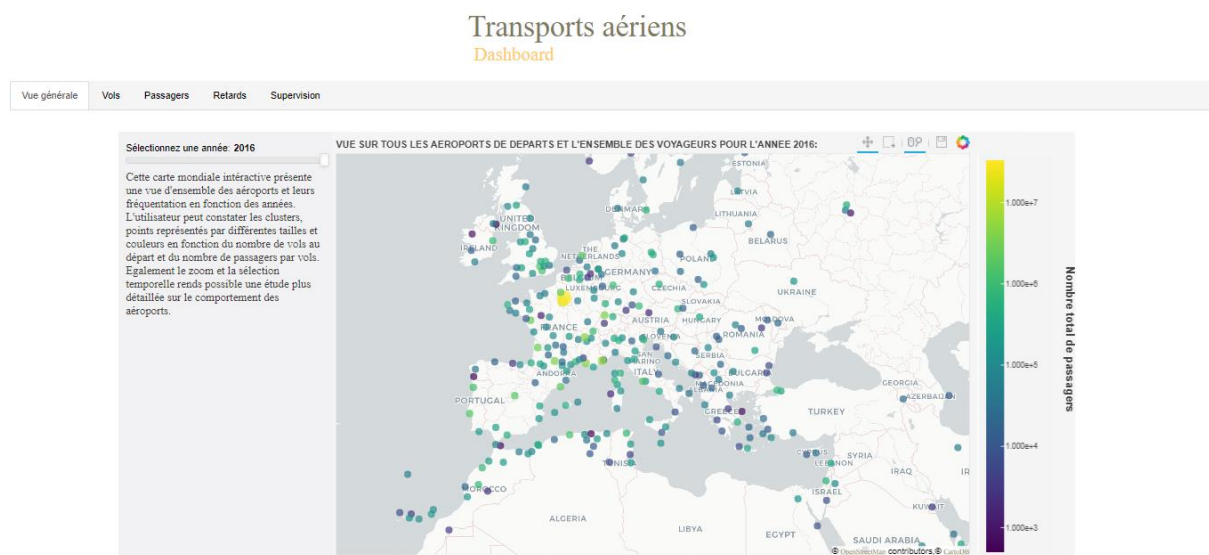
L'architecture de l'application est la suivante : [voir ci-contre]
L'application à l'état actuel est déployable car une étape de refactoring a été effectuée, la rendant modulaire et commandé par un fichier main. L'application est donc maintenable. Le dossier Templates lui contient le HTML et le CSS permettant ainsi une customisation poussée et rapide.

```
main.py
-
datas
-
scripts
  __init__.py
  map_aeroports.py
  passagers.py
  retards.py
  vols.py
-
Templates
  index.html
  styles.css
```

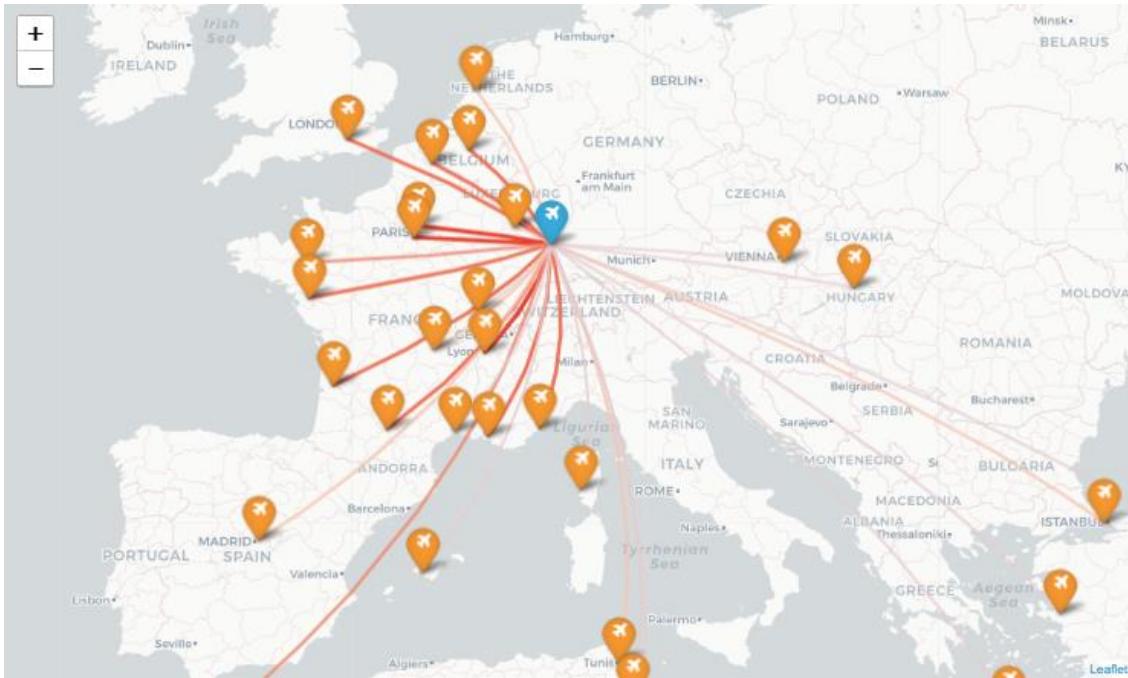
L'application est constituée de plusieurs onglets :

1. Onglet – Vue d'ensemble :

Dans cet onglet, on peut interagir sur une carte du monde. Elle affiche l'évolution du nombre de départs passager par aéroports en fonction des années. L'utilisateur peut interagir avec un slider lui permettant de choisir l'année.

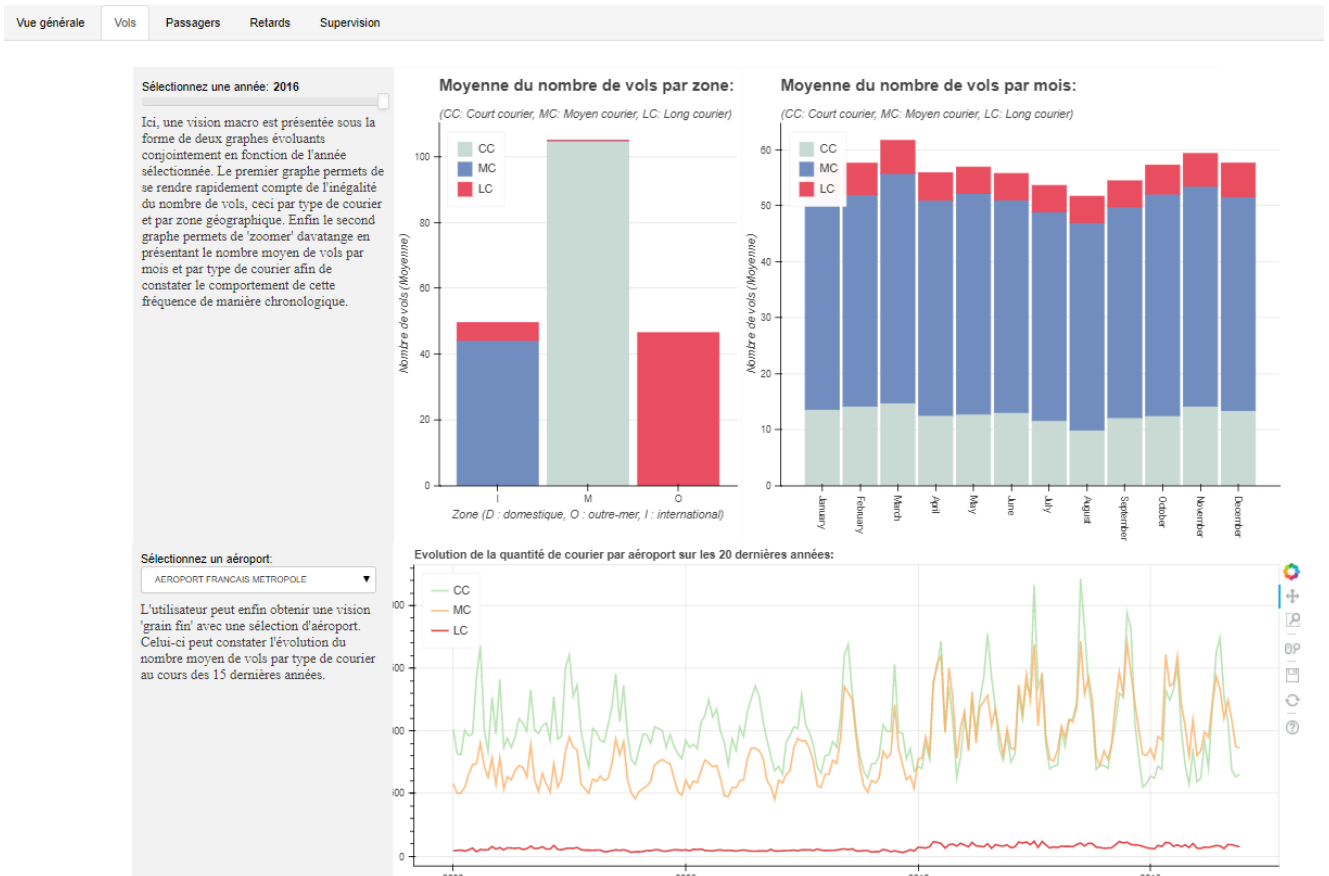


NB: Nous sommes actuellement en cours d'intégration sur Bokeh pour permettre à l'utilisateur de choisir un aéroport et afficher les trajets entre aéroports. L'image ci-dessous montre le résultat que nous avons effectué en utilisant la librairie Folium.



2. Onglet – Vols :

L'idée de cet onglet est de permettre à l'utilisateur d'avoir une vision à la fois macro et « grain fin ». Il a alors accès à trois graphiques interactifs qui se mettent à jours conjointement lors de la sélection d'une année avec le slider. Le premier graphe permet de se rendre rapidement compte de l'inégalité du nombre de vols, ceci par type de courrier et par zone géographique, le second lui affiche le détail par mois.



Pour le troisième graphique (voir vidéo) l'utilisateur peut choisir un aéroport en particulier à l'aide d'un menu déroulant, le graphique se mets à jours de façon automatique, ce dernier affiche la moyenne du nombre de vols en fonction des années et classé par type de courrier.

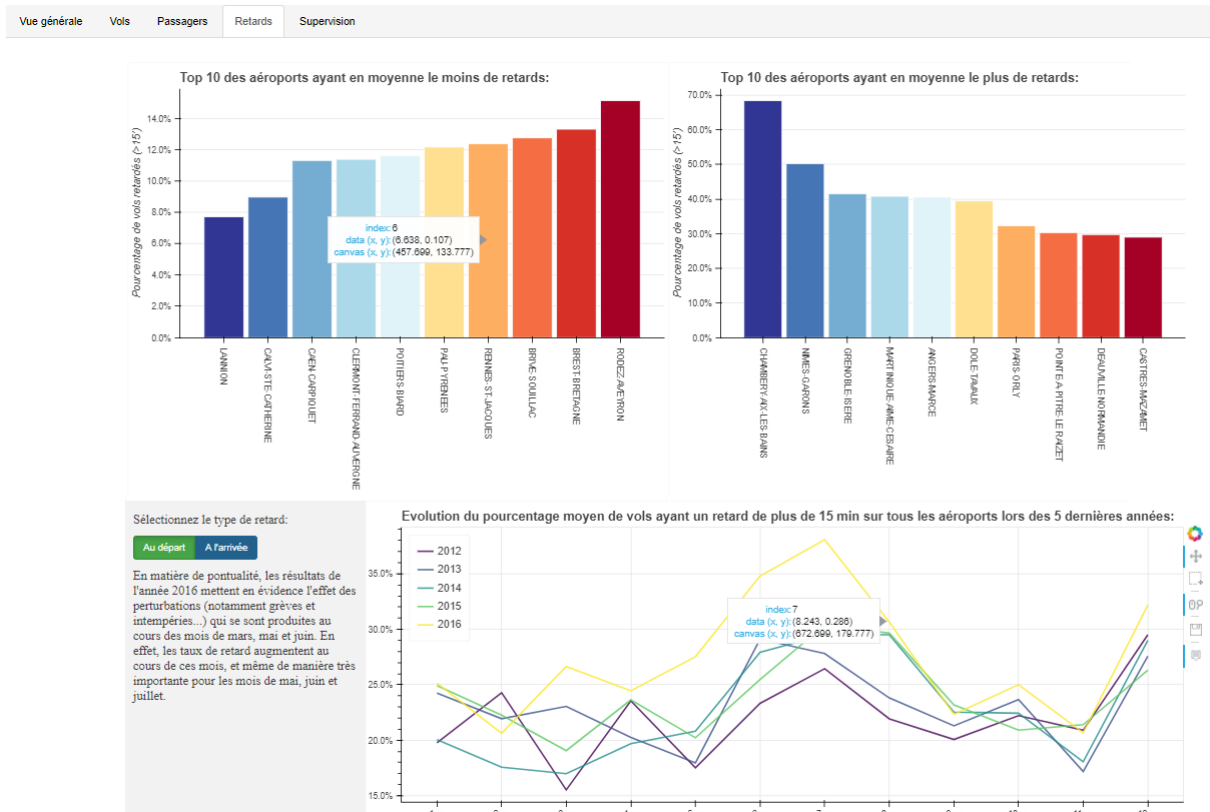
NB : onglet – Passager : Cet onglet suit la même structure que l'onglet précédent ; les graphiques montrent la moyenne des passagers en fonction du temps et classé par type de courrier. Nous sommes actuellement en cours d'intégration sur Bokeh pour fusionner l'onglet Vol avec Passager. L'utilisateur n'aura qu'à choisir à l'aide de bouton :

Select a year: 2016

Nb vols Nb passagers

3. Onglet - Retards :

Dans cet onglet, on a accès à trois graphiques interactifs. Les deux premiers sont les TOP 10 des aéroports ayant en moyenne le moins (resp. le plus) de retards. Le graphique du bas donne l'évolution du pourcentage moyen de retard sur tous les aéroports lors des cinq dernières années. L'utilisateur peut choisir d'observer les retards selon leurs types (au départ, à l'arrivée).



4. Onglet – Consommations :

En cours, intégration. L'idée est d'avoir un graphique dynamique. L'utilisateur pourra observer la consommation de CO2 (ou sélectionner un autre attribue – NOX, COVNM ...) puis le graphique se mettra directement à jours en fonction d'une plage d'année qu'il aura préalablement définie.

5. Onglet – supervision des données :

Dans cet onglet, on affiche sous forme d'histogramme, pour tous les fichiers, le nombre total de données manquantes selon les attributs pour les différents fichiers. Un tableau résumé accompagne le graphique, de façon à avoir un aperçu de la distribution des données. A l'avenir cet onglet pourra comporter des levées d'alertes ainsi qu'un 'admin' pour éditer directement les fichiers CSV.

