

Fouille de Données

Trafic RATP



Réalisé par:

Raphael Trancoso &
Cédric Laguerre

Plan

- Introduction
- Problématique initiale
- Collecte des données
- Statistiques
- Cartographie des données
- Problèmes rencontrés
- Conclusion

1

Introduction

Introduction

- Projet sur les transports en commun
- Importance des statistiques
- Analyse des données



2

Problématique initiale

Problématique initiale

Prédiction des retards de trains

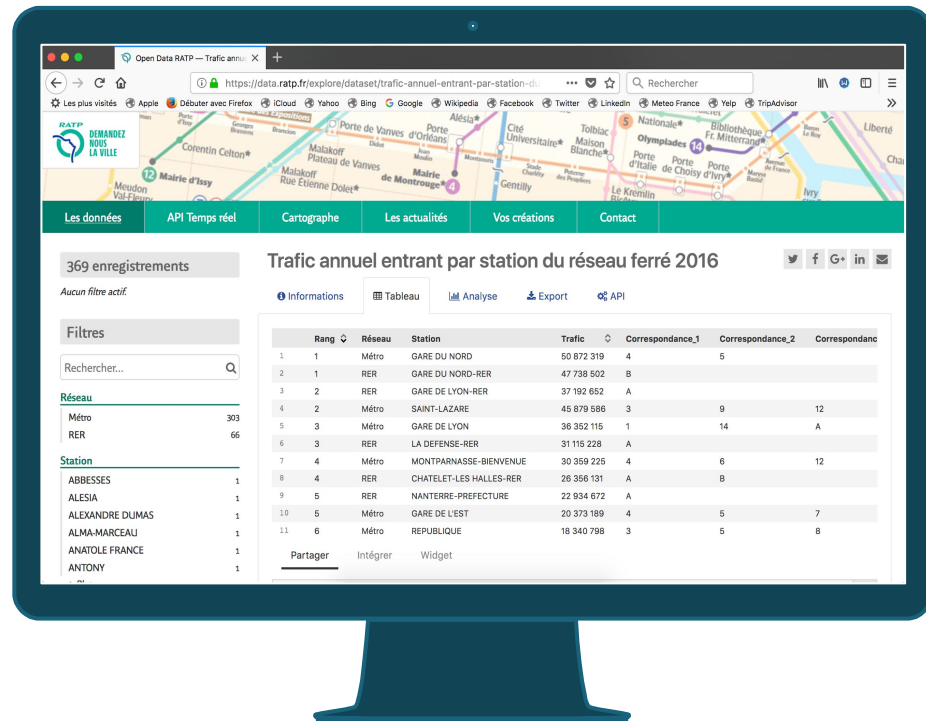


3

Collecte de données

Collecte de données

Site de la RATP



Problématique modifiée

- Données incomplètes
- Nouvelle problématique
- Analyse et représentation du trafic RATP



4

Statistiques



+1,6%

Hausse voyageurs RATP en 2016



3 000 000 000

De personnes transportées en 2016

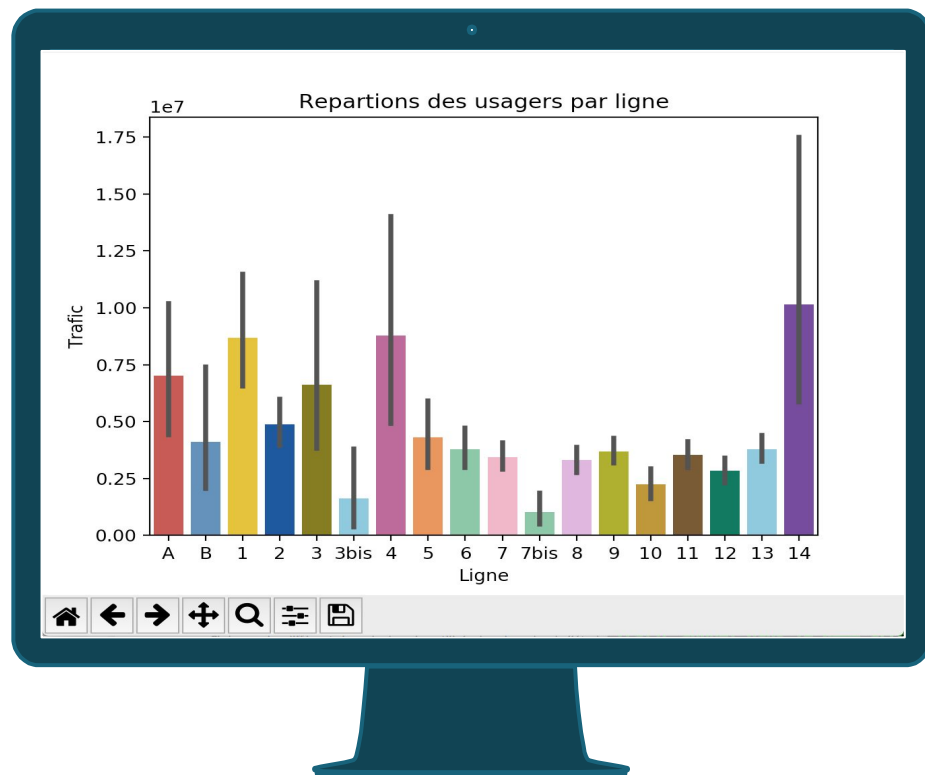


98%

Ponctualité du métro

Graphique représentant le trafic annuel 2016 par ligne

Utilisation de la librairie Seaborn.



5

Cartographie des données

Librairies utilisées

Nuage de points :

- Matplotlib

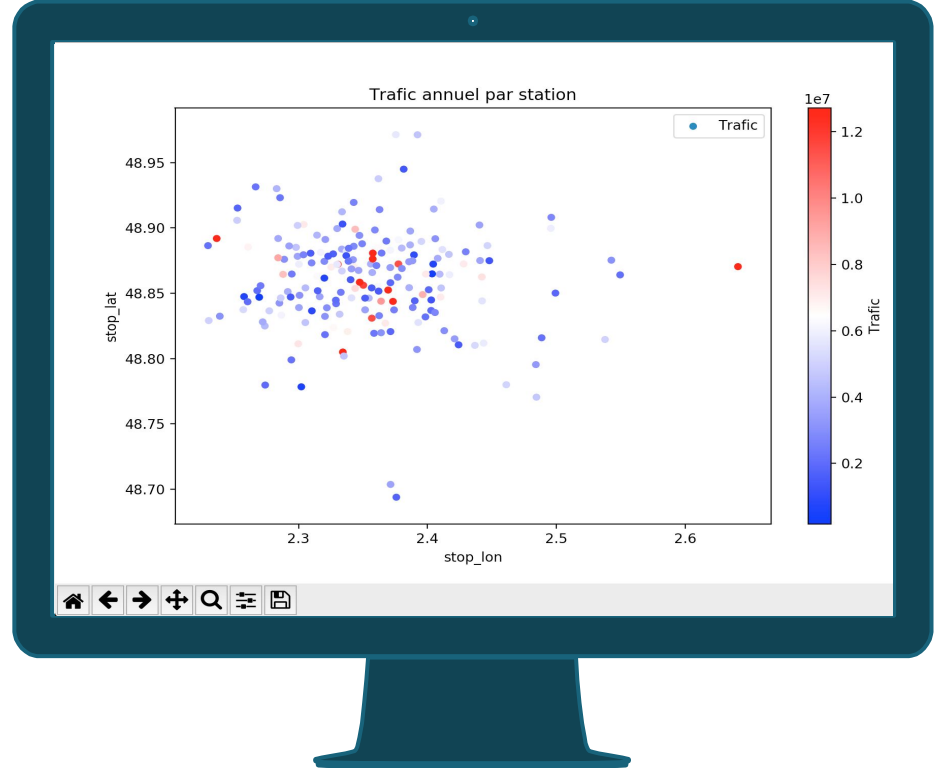
Cartographie :

- Gmplot
- Folium



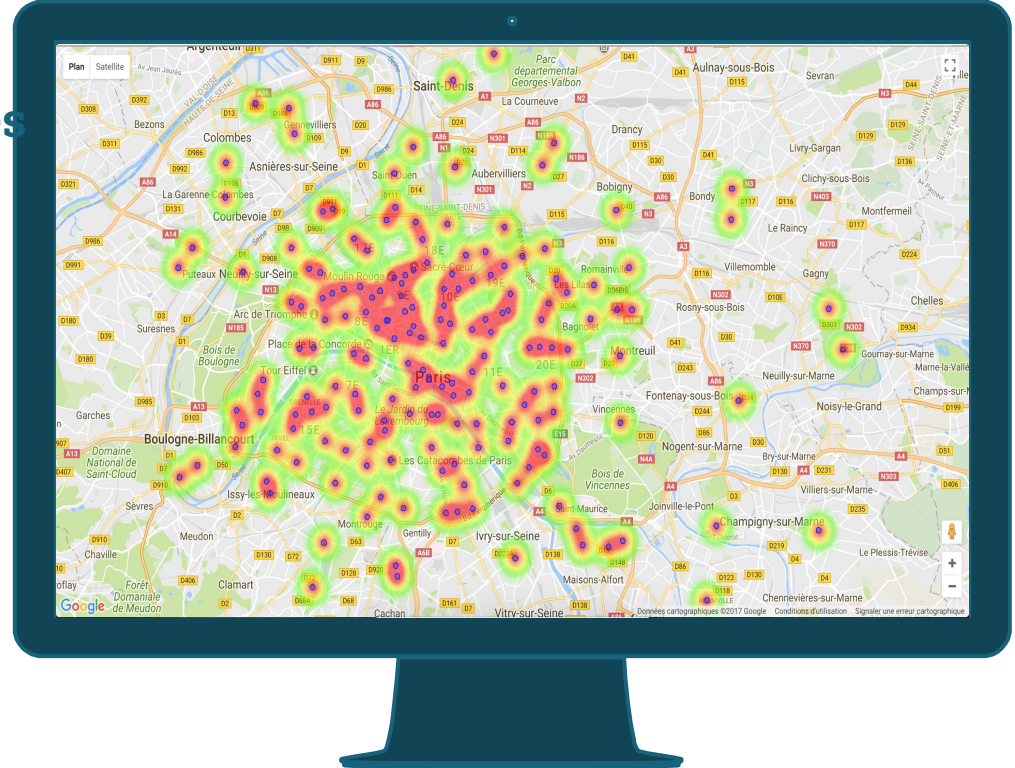
Nuage de points représentant le trafic annuel 2016 par station

- Librairie Matplotlib.
- Utilisation des coordonnées des stations
- Code couleur du trafic



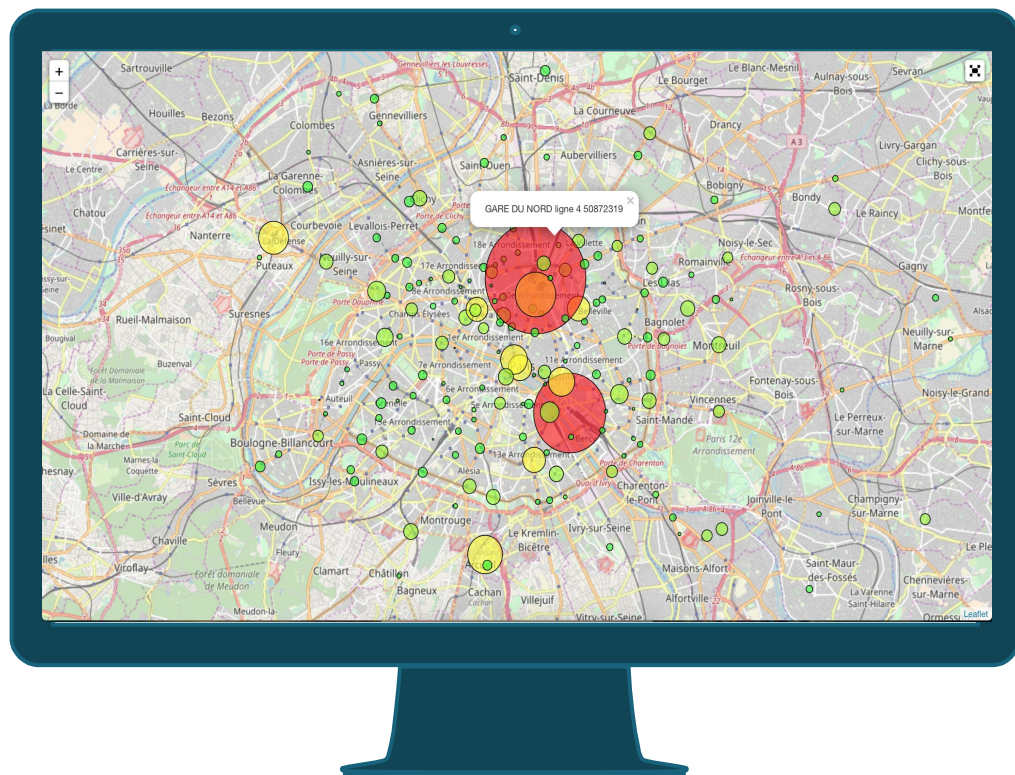
Carte de Paris représentant les stations RATP

- Librairie Gmplot.
- Utilisation des coordonnées des stations
- Heatmap du nombre de stations



Carte de Paris représentant le trafic annuel de chaque station de la RATP

- Librairie Folium.
- Utilisation des coordonnées des stations
- Code couleur du trafic
- Taille des cercles en fonction du trafic



6

Difficultés rencontrées

Difficultés rencontrées

- Données incomplètes
- Représentation cartographique des données

7

Conclusion

Conclusion

- Nouveau langage appris : Python
- Analyse et visualisation satisfaisantes
- Association avec la RATP pour prédire les retards ?

MERCI !

Avez-vous des questions?

Vous pouvez nous contacter par mail :

raph.tran78@gmail.com & ced.laguerre@gmail.com