

PROJET DATAVIZ

BUT SCIENCES DES DONNÉES

PRÉSENTATION DU PROJET





Lancement

- Présentation du sujet et des objectifs
- Présentation des données





Travail par groupe (1 à 4 participants)

- Découverte des données
- Choix de problématique
- Data visualisations



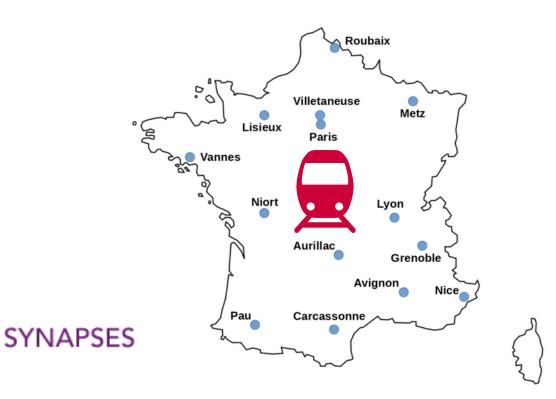


Restitution

Présentation PowerPoint / Dataviz

UN PROJET DE DATAVISUALISTAION

AVEC LES DONNÉES DE LA SNCF



LES AMBITIONS DU PROJET



A partir d'un jeu de données proposé par la SNCF, vous devez réfléchir à une problématique, une question que soulèvent les données.



L'objectif sera de répondre à cette problématique/question par une datavisualisation.



Il s'agit de trouver la ou les bonnes représentations graphiques qui permettent d'illustrer votre problématique.



En fin du projet, chacune des équipes remettra son travail. Vous serez évalués sur :

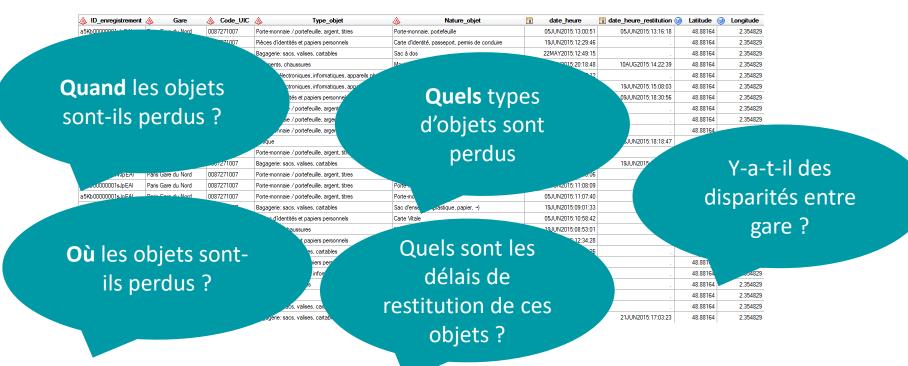
- Le rapport PDF ou PowerPoint avec la présentation du contexte, le déroulement du projet, les choix, les difficultés rencontrées dans l'analyse ...
- Le résultat de la datavisualisation proposée

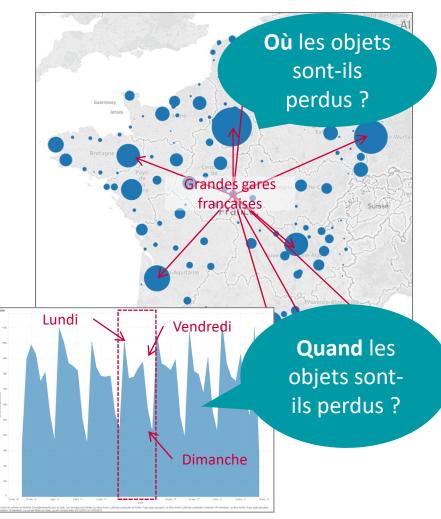


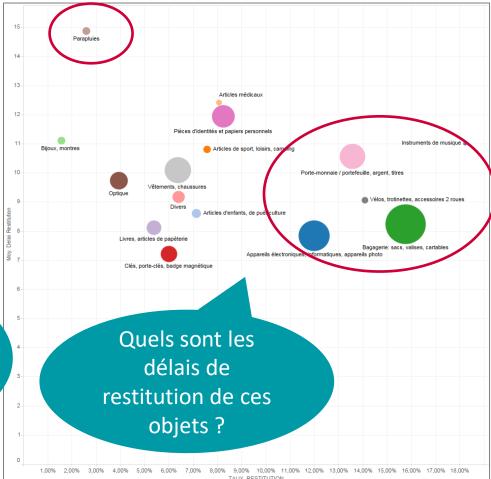
Vous pouvez choisir de jouer un rôle particulier : journaliste, élus locaux, association d'usagers ...



EXEMPLE AVEC LES OBJETS PERDUS







TAUX_RESTITUTION vs. moyenne de Delai Restitution. La couleur affiche des détails associés au/à la Type objet (groupe). La taille correspond au/à la somme de Nombre d'enregistrements. Les repères sont étiquetés par Type objet (groupe).

Présentation du jeu de données

Fréquentation annuelle des gares françaises (montées et descentes) entre 2015 et 2022

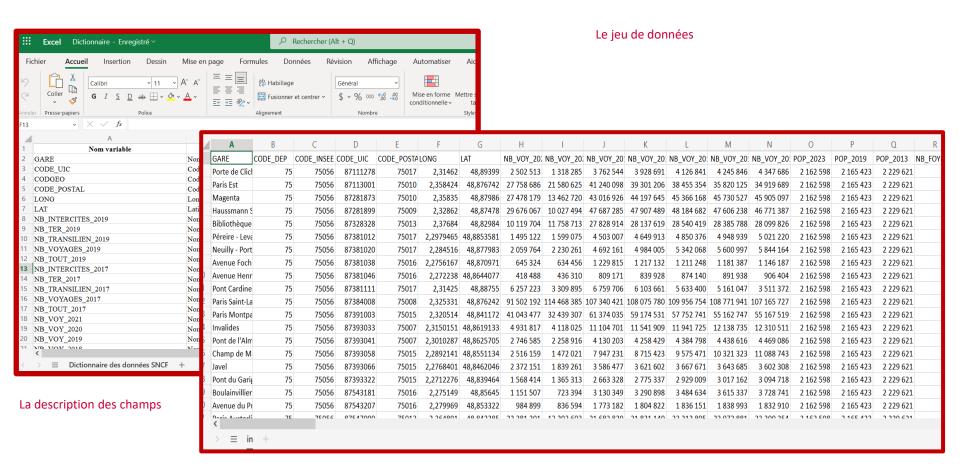
Ces données sont organisées par gare. Elles ont été enrichies avec d'autres sources. Ainsi a-t-on pour chaque gare :

- La position géographique, le code INSEE de la commune correspondante
- L'évolution des offres de train par nature (trains du quotidien TER et Transilien pour l'Ile de France, et trains Longue distance TGV, Intercités pour les lignes classiques) entre 2017 et 2022
- La population de la commune pour 3 années de référence 2013, 2019 et 2023
- Les données immobilières pour la commune (nombre de transactions et prix médian entre 2017 et 2022) pour 2017 sur les 6 derniers mois de l'année et pour 2022 sur les 6 premiers mois de l'année. Attention, ces données ne sont pas disponibles pour certains départements (Haut Rhin , Bas Rhin).
- Des données socio économiques pour l'année de référence 2020 (répartition des revenus par foyer)
- Des informations sur les logements pour l'année de référence 2019

Certaines grandes villes disposent de plusieurs gares (grandes métropoles) pour lesquelles toutes les informations liées à la commune sont dupliquées

Données	Description	Niveau géographique	Historique
Gares	Noms, codes, coordonnées géographiques	Gares de France	
Nombre de départs de trains	Intercités, TER, Transilien, TGV	Gares de France	2017 à 2022
Nombre de voyageurs	Ayant transité dans la gare	Gares de France	2015 à 2022
Population	Nombre d'habitants	Communes (chaque gare est rattachée à sa commune)	2013, 2019, 2023
Revenus	Nombre de foyers dans des tranches de revenus, médiane du niveau de vie	Communes (chaque gare est rattachée à sa commune)	2020
Niveau de vie	Part des ménages fiscaux imposés, taux de pauvreté	Communes (chaque gare est rattachée à sa commune)	2020
Logements	Nombre de logements, résidences principales, secondaires, logements occasionnels, vacants	Communes (chaque gare est rattachée à sa commune)	2019
Emploi	Nombre d'emplois, de chômeurs, de personnes actives, des personnes	Communes (chaque gare est rattachée à sa commune)	2017 à 2022
Immobilier	Nombre de ventes, prix médian au m2	Communes (chaque gare est rattachée à sa commune)	2017 à 2022

LE FICHIER



A vous de jouer!

Quelles sont les données disponibles?

A quelle problématique je souhaite répondre?

Comment puis-je enrichir les données pour répondre à ma problématique?

Quels indicateurs doisje construire pour répondre à ma problématique?

Quelle représentation visuelle serait la plus pertinente pour mettre en évidence mon propos?

Quel client serait intéressé par cette problématique? SNCF? Usagers? Territoire? ...



