

ASSISTANCE AUX PERSONNES HANDICAPÉES DANS LES GARES FRANÇAISES

Objectif	2
Contexte	2
Questions clés.....	2
Données.....	2
Source des données	2
Profil des données	Erreur ! Signet non défini.
Dictionnaire des données	3
Considerations éthiques	4
Confidentialité et anonymat	Erreur ! Signet non défini.
Accessibilité et inclusivité.....	4
Biais et représentation	4
Impact des recommandations.....	4
Utilisation des données publiques	4
Critères d'analyse	Erreur ! Signet non défini.
Exigences du tableau de bord.....	Erreur ! Signet non défini.
Exigences du dépôt Github.....	Erreur ! Signet non défini.

Objective

Construire un tableau de bord interactif présentant visuellement les résultats bien organisés d'une analyse exploratoire avancée réalisée en Python.

Contexte

La **Société nationale des chemins de fer français (SNCF)** est la compagnie ferroviaire nationale française, détenue par l'État.

La SNCF possède plus de 3 000 gares en France, dont 871 ont accueilli des personnes à mobilité réduite entre 2015 et 2022. L'objectif de cette analyse est de proposer des recommandations pour améliorer les conditions d'accueil dans les gares du réseau, et d'identifier les gares prioritaires.

Questions clés

- Existe-t-il des disparités régionales dans la fourniture de services d'assistance au sein du réseau SNCF ?
- Quelle est la corrélation entre le volume de passagers d'une gare et la fréquence de l'assistance fournie ?
- Comment la disponibilité des fauteuils roulants dans les gares affecte-t-elle le nombre et le type d'assistance fournie ?
- Quel impact a la localisation départementale (urbaine vs. rurale) sur le niveau de service fourni aux personnes à mobilité réduite ?
- Sur la base de l'analyse, quelles recommandations actionnables peuvent être faites pour améliorer l'accessibilité et la qualité du service pour les personnes handicapées à travers le réseau SNCF ?

Hypothèses

- Il existe une corrélation positive entre le volume de passagers d'une gare et la fréquence de l'assistance fournie.
- Il existe des disparités régionales dans la fourniture de services d'assistance, certaines zones géographiques recevant plus de services d'assistance que d'autres, ce qui pourrait être influencé par la localisation des gares (urbaines vs. rurales).

Données

Source des données

Ce jeu de données provient du site ressources.data.sncf.com, mais peut également être trouvé sur le site data.gouv.fr.



Profil des données

Ce jeu de données fournit le nombre d'accompagnements pour les personnes à mobilité réduite dans les gares françaises depuis 2015, ainsi que le niveau d'assistance fourni dans les gares et la liste des gares où des fauteuils roulants sont disponibles pour le public.

La donnée:

- Est open-source;
- Provient d'une source authentique (SNCF).
- Incluent des noms de colonnes non anonymisés.
- Datent d'entre 2015 et 2022.
- Contient au moins 5 variables continues.
- Contient au moins 3 variables catégoriques.
- Contient 51,861 lignes.
- Incluent des composantes géographiques.

Dictionnaire des données

accompagnement-pmr-gares.csv ((51,861 lignes, 17 colonnes) :

- **month**: Month the data was recorded
- **station_code**: Station identifier
- **station_name**: Name of the station
- **region_sncf**: SNCF region in which the station is located
- **departement**: Department in which the station is located
- **zip_code**: Postal code of the municipality in which the station is located
- **municipality**: Municipality in which the station is located
- **longitude**: Longitude position of the station
- **latitude**: Latitude position of the station
- **station_category**: 3 categories of stations (3, 2 and 1):
 - Category 3: passenger stations of national interest. These are stations with at least 250,000 passengers a year on national and international passenger services, or where these same passengers account for 100% of the total.
 - Category 2: passenger stations of regional interest. In each region, the management perimeter corresponds to all stations that do not belong to category A, but have a total ridership of at least 100,000 passengers per year.
 - Category 1: passenger stations of local interest. Their management perimeter corresponds, in each region, to all the stations in this category. The fee is set, by region, for all stations in this category.
- **service_level**: 4 levels of service to disabled people (3, 2, 1 and 0):
 - Level 3: Fully independent station access
 - Level 2: Station accessible with the help of an agent
 - Level 1: Station not accessible but alternative transport available
 - Level 0: Station not accessible
- **wheelchair_available_in_station**: Are wheelchair available in the station (1=Yes or 0=No)



- **simple_assistance**: Number of assistances requiring no equipment
- **ramp_assistance**: Number of assistances requiring a ramp
- **wheelchair_assistance**: Number of assistances requiring a wheelchair
- **ramp_and_wheelchair_assistance**: Number of assistances requiring a ramp and a wheelchair •
- **total_assistance**: Total number of assistances

Considerations éthiques

Confidentialité et anonymat

Bien que le jeu de données ne contienne pas d'informations personnellement identifiables (PII) sur les personnes recevant une assistance, il est essentiel de garantir l'anonymat et la confidentialité des individus, en particulier dans les jeux de données qui pourraient être étendus ou combinés avec d'autres sources de données à l'avenir.

Accessibilité et inclusivité

L'analyse doit être menée et présentée de manière à être accessible à un large public, y compris les personnes handicapées. Cela inclut la prise en compte de l'accessibilité du tableau de bord résultant et de toutes les visualisations.

Biais et représentation

Il convient de veiller à identifier et à atténuer tout biais dans les données qui pourrait influencer les conclusions ou les recommandations. Cela inclut la prise en compte de la question de savoir si le jeu de données représente adéquatement la diversité des besoins parmi les personnes à mobilité réduite.

Impact des recommandations

Les recommandations basées sur l'analyse doivent prendre en compte l'impact potentiel sur toutes les parties prenantes, y compris les personnes à mobilité réduite, le personnel des gares et la communauté au sens large. Les considérations éthiques doivent guider les propositions pour s'assurer qu'elles favorisent l'inclusivité, la sécurité et le respect des besoins individuels.

Utilisation des données publiques

Bien que le jeu de données soit open source et provienne d'une source autorisée (SNCF), il est important d'utiliser les données de manière responsable, en respectant leur contexte et les intentions derrière leur mise à disposition publique.

Critères d'analyse

- Analyse exploratoire à travers des visualisations (nuages de points, cartes de corrélation, diagrammes de paires et diagrammes catégoriques)
- Analyse géospatiale à l'aide d'un fichier de formes ;
- Analyse régression ;
- Analyse de cluster ;



- Analyse de séries temporelles
- Récit d'analyse et résultats finaux (présentés dans votre tableau de bord).

Exigences du tableau de bord

- Être conçu avec un cas d'utilisation à l'esprit (répondre aux questions clés) ;
- Être créé sur Tableau Public ;
- Être interactif ;
- Respecter les meilleures pratiques de conception visuelle ;
- Inclure une page d'introduction décrivant le projet (données et objectif)
- Inclure les résultats pertinents de l'analyse exploratoire visuelle initiale ;
- Inclure une explication sur la manière dont les résultats de l'analyse exploratoire ont conduit à la définition des questions de recherche et/ou des hypothèses ;
- Contenir une composante géospatiale ;
- Aborder les questions/hypothèses définies à l'aide de techniques analytiques avancées. Par exemple :
 - Analyse de régression ;
 - Analyse de clusters ;
 - Analyse de séries temporelles ;
- Inclure une page de résumé des résultats expliquant comment les résultats répondent (ou non !) aux questions de recherche/hypothèses initiales ;
- Inclure des détails sur les limites du projet ;
- Inclure une proposition des prochaines étapes pour une analyse plus approfondie.

Exigences pour le dépôt GitHub

- Code en Python ;
- Structure de dossiers logique ;
- Dossiers et fichiers respectant les conventions de nommage standard de l'industrie
- Scripts Jupyter prêts à être ajoutés à un portfolio pour chaque étape (avec des commentaires dans le code, une structure organisée et un code propre et fonctionnel)
- Un fichier README contenant:
 - Une description du projet:
 - Des détails sur la ou les sources de données:
 - Les questions de recherche:
 - Les procédures de nettoyage des données ;
- Un lien vers le tableau de bord Tableau.

