ASSISTANCE FOR DISABLED PEOPLE IN FRENCH STATIONS

Joris Jan

Objective

To build an interactive dashboard visually showcasing well-curated results of an advanced exploratory analysis conducted in Python.

Contexte

The **Société nationale des chemins de fer français** (abbreviated as **SNCF**; French for "National Company of the French Railways") is France's national state-owned railway company.
SNCF owns more than 3,000 stations in France, 871 of which welcomed people with reduced mobility between 2015 and 2022. The aim of this analysis is to propose recommendations for improving reception conditions in the network's stations, and to identify priority stations.

Key Questions

- Are there any regional disparities in the provision of assistance services within the SNCF network?
- What is the correlation between the station's passenger volume and the frequency of assistance provided?
- How does the availability of wheelchairs in stations affect the number and type of assistance provided?
- What impact does the departmental location (urban vs. rural) have on the level of service provided to people with reduced mobility?
- Based on the analysis, what actionable recommendations can be made to improve the accessibility and quality of service for disabled individuals across the SNCF network?

Hypotheses

- There is a positive correlation between the station's passenger volume and the frequency of assistance provided.
- There are regional disparities in the provision of assistance services, with certain geographic areas receiving more assistance services than others, which could be influenced by urban vs. rural station locations.

Data

Data Source
This dataset comes from the ressources.data.sncf.com site, but can also be found on the data.gouv.fr site (the single interministerial portal designed to gather and make freely available all public information from the French State, its public administrative establishments, local authorities and public or private entities entrusted with a public service mission).

Data Profile

This dataset provides the number of escorts for people with reduced mobility in French stations since 2015, as well as the level of assistance provided in stations and the list of stations where wheelchairs are available to the public.

The data:
- Is open source;
- Come from an authentic/authoritative source (SNCF);
- Include non-anonymized column names;
- Is no more than 3 years old (from 2015 to 2022);
- Contain at least 5 continuous variables (excluding index or ID variables, dates, years, etc.);
- Contain at least 3 categorical variables (excluding index or ID variables, dates, years, etc.);
- Contain 51,619 rows;
- Include geographical components

Data Dictionary

assistance for disabled people in stations.csv (51,619 rows, 17 columns):
- 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

Ethical Considerations

Privacy and Anonymity

Although the dataset does not include personally identifiable information (PII) about individuals receiving assistance, ensuring the anonymity and privacy of individuals is essential, especially in datasets that might be expanded or combined with other data sources in the future.

Accessibility and Inclusivity

The analysis should be conducted and presented in a manner that is accessible to a wide audience, including those with disabilities. This includes considering the accessibility of the resulting dashboard and any visualizations.

Bias and Representation

Care should be taken to identify and mitigate any biases in the data that could influence findings or recommendations. This includes considering whether the dataset adequately represents the diversity of needs among people with reduced mobility.

Impact of Recommendations

Recommendations based on the analysis should consider the potential impact on all stakeholders, including people with reduced mobility, station staff, and the broader community. Ethical considerations should guide proposals to ensure they promote inclusivity, safety, and respect for individual needs.

Use of Public Data

While the dataset is open source and from an authoritative source (SNCF), it's important to use the data responsibly, respecting its context and the intentions behind making it publicly available.

Analysis Criteria

- Exploratory analysis through visualizations (scatterplots, correlation heatmaps, pair plots, and categorical plots);
- Geospatial analysis using a shapefile;
- Regression analysis;
- Cluster analysis;
- Time-series analysis;
- Analysis narrative and final results (presented in your dashboard).

Dashboard Requirements

- Be designed with a use case in mind (answering key guidance questions);
- Be created in Tableau Public;
- Be interactive;
- Adhere to visual design best practices;
- Include an introduction page that describes the project (data and purpose);
- Include relevant result(s) of initial visual exploratory analysis;
- Include an explanation for how the results of the exploratory analysis resulted in defining
- research questions and/or hypotheses;
- Contain a geospatial component;
- Address the defined questions/hypotheses using advanced analytical techniques. For example:
  - Regression analysis;
  - Cluster analysis;
  - Time-series analysis.
- Include a results summary page explaining how the results do (or don't!) address the initial research questions/hypotheses;
- Include details on the limitations of the project;
- Include a proposal of the next steps for further analysis.

GitHub Repository Requirements

- Python code;
- Logical folder structure;
- Folders and files that follow industry-standard naming conventions;
- Portfolio-ready Jupyter scripts for every steps (complete with code comments, organized structure, and clean, functioning code);
- A README file containing:
  - A description of the project;
  - Details of the data source(s);
  - Research questions;
  - Cleaning procedures;
  - A link to the Tableau dashboard.