

Public Transport Efficiency Analysis Phase-2

2023-10-11

```
matplotlib inline

import numpy as np # linear algebra
import pandas as pd # data processing, CSV
file I/O (e.g. pd.read_csv)

import matplotlib.pyplot as plt

import datetime

import os

from math import sqrt

import warnings


## For Multiple Output in single cell
from IPython.core.interactiveshell import
InteractiveShell

InteractiveShell.ast_node_interactivity =
'all'

warnings.filterwarnings('ignore')

data = pd.read_csv('../input/unisys/
ptsboardingsummary/20140711.CSV')

data.shape

data.head(10)
```

2023-10-11

Analyzing the efficiency of public transport in the future will likely involve several key areas of focus:

Data Analytics: Utilizing advanced data analytics and machine learning to process vast amounts of transportation data, including real-time information, to optimize routes, schedules, and resource allocation.

Sustainability: Assessing the environmental impact and sustainability of public transport systems, with an emphasis on reducing emissions and energy consumption through electrification and eco-friendly technologies.

Autonomous Vehicles: Evaluating the integration of autonomous vehicles into public transport networks, considering safety, cost-effectiveness, and passenger experience.

Smart Infrastructure: Developing smart infrastructure that can communicate with vehicles and passengers to enhance efficiency, safety, and convenience.

Multi-Modal Integration: Promoting seamless integration of various modes of public transport, such as buses, trains, trams, and shared mobility services, to create a comprehensive transportation network.

Accessibility: Ensuring public transport remains accessible to all, including people with disabilities and those in underserved areas, through inclusive design and technology.

Ticketing and Payment Systems: Enhancing ticketing and payment systems, potentially moving toward contactless and mobile payment solutions to streamline the boarding process.

User Experience: Improving the overall passenger experience, including real-time updates, wayfinding, and comfort during the journey.

Environmental Impact Assessment: Continuously monitoring and evaluating the environmental impact of public transport systems, and seeking ways to minimize negative effects.

Economic Viability: Assessing the financial sustainability of public transport systems, exploring innovative funding models and partnerships to support long-term operation.

Community Engagement: Involving the community in decision-making processes, seeking feedback, and tailoring public transport services to meet local needs.

Security and Safety: Implementing advanced security measures and safety protocols to protect both passengers and infrastructure from various threats.

Resilience: Preparing public transport systems to withstand and recover from natural disasters and other disruptive events.

Policy and Regulation: Continuously adapting policies and regulations to accommodate evolving technologies and ensure fair competition in the transport sector.

Environmental Goals: Aligning public transport efficiency analysis with broader environmental and climate goals, such as reducing carbon emissions and promoting sustainable urban development.

As technology and urban landscapes evolve, the analysis of public transport efficiency will require ongoing adaptation and innovation to meet the changing needs of communities and the

Predictive modelling of public transport efficiency analysis

Predictive modeling for public transport efficiency analysis involves using historical data and various statistical or machine learning techniques to forecast future performance and identify areas for improvement. Here's a high-level overview of the process:

Data Collection: Gather relevant data, including information on routes, schedules, passenger counts, weather conditions, and any other factors that may affect public transport efficiency.

Data Preprocessing: Clean and preprocess the data, handling missing values, outliers, and formatting issues.

Feature Selection: Identify the most relevant features (variables) that are likely to impact efficiency, such as traffic patterns, maintenance schedules, and demand fluctuations.

Model Selection: Choose an appropriate predictive modeling technique, which could include linear regression, time series analysis, decision trees, random forests, or neural networks, depending on the nature of the data and the problem.

Model Training: Use historical data to train the chosen model, adjusting its parameters to minimize prediction errors.

Validation and Testing: Evaluate the model's performance using validation data and testing data to ensure it can make accurate predictions on new, unseen data.

Prediction: Use the trained model to make predictions about future public transport efficiency. This could involve forecasting delays, optimizing schedules, or predicting passenger demand.

Interpretation: Analyze the model's results to gain insights into the factors affecting efficiency. Identify areas for improvement and potential strategies to enhance public transport services.

Interpretation: Analyze the model's results to gain insights into the factors affecting efficiency. Identify areas for improvement and potential strategies to enhance public transport services.

Deployment: Implement the findings and recommendations from the predictive model in the real-world public transport system. This may involve adjusting schedules, routes, or maintenance procedures.

Continuous Monitoring: Continuously monitor the system's performance and update the model as new data becomes available to ensure ongoing improvements in public transport efficiency.

Remember that the success of predictive modeling for public transport efficiency analysis depends on the quality of data, the choice of modeling techniques, and the commitment to continuous improvement based on the model's insights.

2023-10-10

Data exploration for public transport efficiency analysis involves examining and understanding the dataset related to public transportation to gain insights. Here are some steps you can take:

Data Collection: Gather relevant data sources, which may include information on routes, schedules, ridership, vehicle types, and more. This data may come from public transportation authorities, government agencies, or other sources.

Data Cleaning: Clean the data to remove duplicates, missing values, and inconsistencies. Ensure that the data is in a structured format for analysis.

Descriptive Statistics: Calculate basic statistics such as mean, median, standard deviation, and range for relevant variables. This helps in getting a sense of the data's distribution.

Data Visualization: Create visualizations like histograms, scatter plots, and box plots to visualize the distribution and relationships within the data. This can help identify trends or outliers.

Geospatial Analysis: If your data includes geographical information, use maps and spatial analysis tools to explore patterns in public transport routes and efficiency across different areas.

Time-Series Analysis: If time-related data is available, analyze trends over time, seasonality, and patterns in ridership or performance.

Correlation Analysis: Determine if there are correlations between different variables, such as ridership and fuel consumption, which could impact efficiency.

Cluster Analysis: Group similar routes or regions together based on specific characteristics, which can help in identifying areas for improvement.

which could impact efficiency.

Cluster Analysis: Group similar routes or regions together based on specific characteristics, which can help in identifying areas for improvement.

Machine Learning: Consider using machine learning techniques to build predictive models for public transport efficiency based on historical data.

Feedback Loops: If possible, incorporate feedback mechanisms to continually improve public transport efficiency based on the insights gained from data analysis.

Remember that the specific analysis methods will depend on the nature of the data and the goals of your public transport efficiency analysis. The insights obtained through data exploration can guide decisions and improvements in the public transportation system.

Data aggregation of public transport efficiency analysis

Data aggregation for public transport efficiency analysis involves collecting and summarizing relevant data to assess the performance of public transportation systems. Here are the steps to perform data aggregation for this analysis:

Data Sources:

Identify the sources of data, which may include public transit agencies, government reports, transportation apps, and surveys.

Data Collection:

Collect data on various aspects of public transport, such as ridership, routes, schedules, vehicle types, ticket prices, and maintenance records.

Data Cleaning:

Clean and preprocess the data to remove duplicates, missing values, and inconsistencies.

Data Integration:

Combine data from different sources into a single dataset, ensuring that data from various agencies and formats are compatible.

Data Transformation:

Convert data into a standardized format for analysis. For example, convert timestamps into a common time zone and currency values into a common currency.

Aggregation Metrics:

Define the key performance indicators (KPIs) and metrics that will be used to assess public transport efficiency. Common metrics include on-time performance, average travel time, passenger load factors, and cost per passenger.

Temporal Analysis:



External features:

Transportation Authorities: Public transportation agencies often provide data on routes, schedules, ridership, and performance metrics. You can contact local or regional transportation authorities to request access to their data.

Open Data Portals: Many cities and regions have open data portals that offer public access to transportation-related datasets. These portals may include information on bus and train schedules, ridership numbers, and more.

GPS and IoT Sensors: Some public transportation systems use GPS and Internet of Things (IoT) sensors to track the movement of vehicles in real-time. This data can be valuable for analyzing efficiency.

Surveys and Passenger Feedback: Surveys and feedback from passengers can provide insights into the quality of public transport services, including factors like punctuality, cleanliness, and safety.



Mobile Apps and Ticketing Systems: Mobile apps and ticketing systems used by public transport riders can generate data on passenger behavior, such as boarding and alighting locations, payment methods, and journey durations.

Academic Research: Academic studies and research papers may contain valuable data and analyses related to public transport efficiency. Libraries and academic databases can be good sources for such information.

Government Reports: Government agencies often publish reports and studies on public transport performance and efficiency. These reports may contain valuable data and insights.

Crowdsourced Data: Platforms like Google Maps, Waze, and various transit apps collect data from users and can provide information on real-time traffic conditions, delays, and route recommendations.