10Alytics

OPTIMIZING PUBLIC TRANSIT OPERATIONS

An Exploratory Data Analysis of public transportation

10Alytics

-- Business Introduction --

MetroMove Transit Solutions is a public transportation service provider operating in multiple cities.

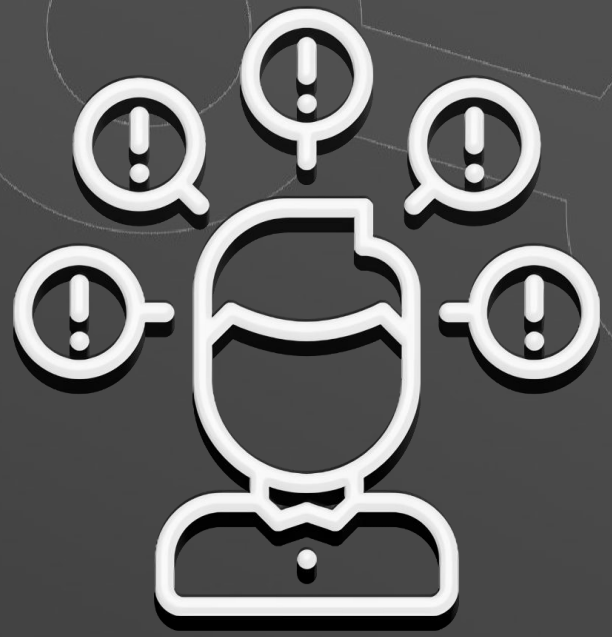
They manage and analyze thousands of daily trips taken via buses, trains, ferries, and trams.

The company's mission is to provide **efficient, affordable, and timely public transportation** services while leveraging data to **improve passenger experience** and **optimize operations.**

10Alytics

-- PROBLEM STATEMENT --

MetroMove has collected **a large volume of trip data** but lacks insights into **trip performance, passenger behavior, and fare patterns** due to **messy, inconsistent, and incomplete records**. As part of their new data-driven initiative, they want to **clean, explore, and summarize** their trip records to identify inefficiencies and patterns that can drive operational improvements.



10Alytics

-- Rationale for the Project—

- **Effective decision-making relies on**
  - Understanding passenger usage patterns.
  - Evaluating the performance of different transport modes.
  - Analyzing how trip characteristics impact customer experience.
- **This project provides a real-world simulation by**
  - Requiring cleaning and preprocessing of raw data.
  - Encouraging pattern discovery through exploratory data analysis.
  - Developing skills in communicating data-driven insights.
  - Working within the context of a transportation-focused dataset.

10Alytics

--Deliverables—

Data Cleaning

Data Explorations

Descriptive Statiscs

Univariate/Bivariate/
Multivariate Analysis

Visualizations

Jupyter Notebook
containing comments
and recommendations

10Alytics

-- Data Description --

Column	Description
Trip_ID	Unique identifier for each trip
Mode_of_Transport	Type of transport used: Bus, Train, Ferry, or Tram (includes inconsistencies)
Departure_Station	Station where the trip starts (contains whitespace errors)
Arrival_Station	Station where the trip ends (inconsistent casing)
Departure_Time	Exact date and time when the trip departed
Passenger_Count	Number of passengers on the trip (includes missing values)
Fare_Amount	Amount paid by the passengers for the trip (includes missing values)
Trip_Duration_Minutes	Duration of the trip in minutes (includes missing values)
Trip_Date	Date on which the trip occurred
Day_of_Week	Day of the week on which the trip occurred