**Project Name:** Public Transportation Efficiency Analysis

**Date:** 26/10/2023

**Project Team Member:** Varsha, Abinaya, Tharani, Neha Shalini, Vinoth

**1. Data Sources:**

Gather data from various sources, including GPS tracking, ticketing systems, passenger counts, maintenance logs, and more.

Integrate data from different modes of transportation (buses, trams, subways, etc.) into a unified system.

**2. Key Performance Indicators (KPIs):**

Define the KPIs that are essential for monitoring efficiency. These might include:

On-time performance

Ridership and occupancy rates

Route punctuality

Vehicle maintenance status

Fuel or energy consumption

Revenue and cost metrics

Customer satisfaction ratings

**3. Visualization Tools:**

Select appropriate data visualization tools and technologies. Common choices include:

Real-time maps showing vehicle locations and routes.

Line charts for tracking on-time performance over time.

Bar charts for comparing ridership and occupancy rates.

Gauges or meters to display vehicle maintenance status.

Cost and revenue breakdowns through pie charts.

**4. Dashboard Design:** Design an intuitive and user-friendly dashboard layout.

Prioritize the most critical KPIs for quick and easy access.

Use color coding to indicate status (e.g., green for good, red for issues).

Include filters and drill-down options for more detailed insights.

**5. Real-Time Updates:** Ensure that the dashboard provides real-time or near-real-time updates.Set up alerts for critical issues or deviations from established benchmarks.

**6. Geographic Visualization:** Include a map view that displays the current locations of all vehicles.Use color coding and markers to indicate vehicle status and routes.

**7. Historical Data Analysis:** Incorporate historical data for trend analysis.

Show performance trends over days, weeks, or months.

**8. User Access Control:** Implement user access control to ensure that only authorized personnel can view or modify sensitive data.

**9. Mobile Compatibility:**

Ensure that the dashboard is mobile-friendly for access on the go.

**10. Data Analytics Integration:**

Connect the dashboard to data analytics tools to provide deeper insights and predictive analysis.

**11. Reporting and Exporting:**

Allow users to generate reports and export data for further analysis or sharing.

**12. Feedback Mechanism:** - Include a feedback mechanism to collect input from drivers, staff, and passengers.

**13. Continuous Improvement:** Regularly update the dashboard based on user feedback and evolving transportation needs.

**14. Public Outreach:** Consider making certain data, such as service disruptions or delays, available to the public to enhance transparency.

**15. Training and Support:** Provide training and support to personnel who will be using the dashboard.

A well-designed public transport efficiency dashboard can be a powerful tool for monitoring, managing, and improving the performance of public transportation systems. It enables data-driven decision-making, which can lead to cost savings, improved service quality, and increased public satisfaction with the transportation network.

**Codings for Python:**

import pandas as pd

import matplotlib.pyplot as plt

dataset\_path = "public\_transport\_data.csv"

df = pd.read\_csv(dataset\_path)

missing\_values = df.isnull().sum()

print("Missing Values:")

print(missing\_values)

summary\_stats = df.describe()

print("\nSummary Statistics:")

print(summary\_stats)

passenger\_counts = df.groupby('Route')['PassengerCount'].sum()

passenger\_counts.plot(kind='bar')

plt.title("Total Passenger Counts by Route")

plt.xlabel("Route")

plt.ylabel("Passenger Count")

plt.show()

df = df.drop\_duplicates()

df['AverageWaitingTime'] = df['WaitingTime (min)'].mean()

on\_time\_percentage = (df['OnTimeArrivals'].sum() / len(df)) \* 100

print("\nOn-Time Performance Percentage:", on\_time\_percentage, "%")

df.to\_csv("cleaned\_public\_transport\_data.csv", index=False)

**Output:**

Date,Route,Distance (km),PassengerCount,OnTimeArrivals,WaitingTime (min)

2023-01-01,RouteA,10,50,45,10

2023-01-02,RouteA,10,55,48,12

2023-01-03,RouteB,15,75,70,15

2023-01-04,RouteB,15,70,63,14

2023-01-05,RouteA,10,52,47,11

**Missing Values:**

Date 0

Route 0

Distance (km) 0

PassengerCount 0

OnTimeArrivals 0

WaitingTime (min) 0

dtype: int64

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Missing Values: | |  |  |  |  |
| Date 0 | |  |  |  |  |
| Route 0 | |  |  |  |  |
| Distance (km) 0 | |  |  |  |  |
| PassengerCount 0 | |  |  |  |  |
| OnTimeArrivals 0 | |  |  |  |  |
| WaitingTime (min) 0 | | |  |  |  |
| dtype: int64 | |  |  |  |  |
|  |  |  |  |  |  |
| Summary Statistics: | |  |  |  |  |
| Distance (km) PassengerCount OnTimeArrivals WaitingTime (min) | | | | | |
| count 5.00000 5.000000 5.000000 5.00000 | | | | |  |
| mean 12.00000 60.400000 54.600000 12.40000 | | | | |  |
| std 2.23607 11.852249 9.629647 2.30217 | | | | |  |
| min 10.00000 50.000000 45.000000 10.00000 | | | | |  |
| 25% 10.00000 52.000000 47.000000 11.00000 | | | | |  |
| 50% 15.00000 55.000000 48.000000 12.00000 | | | | |  |
| 75% 15.00000 65.000000 63.000000 14.00000 | | | | |  |
| max 15.00000 75.000000 70.000000 15.00000 | | | | |  |
|  |  |  |  |  |  |
| On-Time Performance Percentage: 91.66666666666666 % | | | | |  |

