Technical Report: Analysis and Forecasting of Public Transport Data

---

Objective

The goal of this analysis is to explore trends in passenger journeys across five transport services (Local Route, Light Rail, Peak Service, Rapid Route, and School), derive key insights, and forecast passenger counts for the next seven days using a suitable algorithm.

Chosen Algorithm

Moving Average Method

The **Moving Average Method** was chosen for forecasting due to its simplicity, effectiveness for time-series data, and ability to smooth out fluctuations in passenger counts.

Algorithm Description

The **Moving Average Method** calculates the average of data points within a fixed sliding window (e.g., 7 days). This average is used as a forecast for future data points. It's particularly effective for identifying trends and seasonality in datasets with consistent patterns.

- x_i: Passenger count for day i.
- n: Window size (e.g., 7 days).

Model Parameters

1. **Window Size (n)**:
   - A **7-day window** was used to capture weekly trends in passenger journeys. This ensures the forecast accounts for typical variations in travel patterns, such as weekday vs. weekend trends.

2. **Input Data**:
   - Daily passenger counts for each service type (e.g., Local Route, Light Rail) were used to compute the moving average.

3. **Prediction Horizon**:
   - Forecasts were generated for the **next 7 days** for each service type.

Advantages of the Algorithm

- Simple to implement and interpret.
- Reduces noise in the dataset to highlight underlying trends.
- Effective for short-term forecasts in stable data.

Insights from the Dataset

1. **Service Popularity**:
   o Local Route has the highest average passenger count, while School services have the lowest.

2. **Trends**:
   o Passenger journeys show consistent daily variations, with peaks on weekdays and lower counts on weekends.

3. **Busiest and Least Busy Days**:
   o The busiest day had **X passengers** (date: **DD/MM/YYYY**), while the least busy day had **Y passengers** (date: **DD/MM/YYYY**).

4. **Forecast Results**:
   o Predicted counts for Local Route in the next 7 days range between **X1** and **X2** passengers daily.

Limitations

- The Moving Average Method does not account for external factors like holidays or disruptions.
- It is less accurate for datasets with irregular or highly seasonal patterns.

Conclusion

The Moving Average Method is an ideal choice for forecasting passenger trends in this dataset due to its simplicity and effectiveness. The forecast provides actionable insights for planning resources across various transport services.