Public Transport Passenger Journey Analysis & Forecasting

1. Dataset Overview

The dataset consists of 1,918 entries and 7 columns, representing daily public transport passenger journeys for different service types. The columns include:

- Date: Daily data, used as the time index for analysis.

- Local Route: Passenger count for local bus routes.

- Light Rail: Passenger count for light rail services.

- Peak Service: Passenger count for peak hour services.

- Rapid Route: Passenger count for rapid transit.

- School: Passenger count for school-specific routes.

- Other: Additional services with potential missing values.

2. Initial Exploratory Data Analysis (EDA)

We performed some initial data exploration to understand the overall trends and patterns in the data. Key observations include:

- **Summary Statistics**: We see significant variations in passenger counts across service types. For instance, the average daily journeys for 'Rapid Route' and 'Local Route' are the highest, while 'Peak Service' has a lower average count.
- **Trends Over Time**: Line plots reveal seasonal patterns in daily passenger journeys, particularly for Local Route, Rapid Route, and School services.
- **Missing Data**: Missing values are handled using forward-fill to maintain data continuity.

3. Potential Business Insights & Predictive Opportunities

Here are five key business insights and predictive opportunities derived from the dataset:

1. **Demand Forecasting for Resource Allocation**:

   Insight: Identify peak travel times for different services.

   Prediction: Forecast daily or weekly demand to optimize staffing and vehicle allocation.

2. **Service Type Popularity & Infrastructure Investment**:

   Insight: Analyze which service types are most utilized and which are underperforming.

   Prediction: Use trend analysis to recommend infrastructure investments for high-demand services.

3. **Seasonal Analysis for Promotional Planning**:

   Insight: Determine seasonal usage patterns (e.g., school route spikes).

   Prediction: Create promotional campaigns targeting high-usage periods or encouraging off-peak travel.

4. **Passenger Load Optimization**:

   Insight: Identify underutilized routes or time periods.

   Prediction: Suggest potential changes to service schedules to improve efficiency and reduce costs.

5. **Performance Tracking for Rapid Route**:

   Insight: Compare daily trends in Rapid Route services to assess reliability.

   Prediction: Develop a model to predict demand surges and proactively adjust services to maintain quality.