Key insights from the datasets

- Overall Passenger Volume Trend
- Weekly Seasonality and Day-of-Week Effects
- Service-Specific Trends
- School Service Seasonality
- Peak Service Usage Patterns

Technical Report on Forecasting Algorithm

Algorithm Used:

SARIMA (Seasonal AutoRegressive Integrated Moving Average)

Why SARIMA?

The passenger data shows clear trends and weekly seasonal patterns. SARIMA is well-suited to handle this kind of time series with seasonality, especially daily data with a 7-day cycle. It is a popular and interpretable method commonly used in transport forecasting.

Methodology

- Data Preparation:
  Cleaned the dataset, converted the Date column to datetime, set it as index, and filled missing dates by carrying forward passenger counts. Grouped data by date and service type.
- Exploratory Analysis:
  Examined trends and weekly patterns using plots to understand passenger behavior over time.
- Model Building:
  Chose SARIMA parameters: non-seasonal order (1,1,1) and seasonal order (1,1,0,7) reflecting weekly seasonality. Trained on all data except the last 7 days, then forecasted the next 7 days for each service.
- Evaluation:
  Measured accuracy with RMSE and compared predicted values against actual counts through visualization.