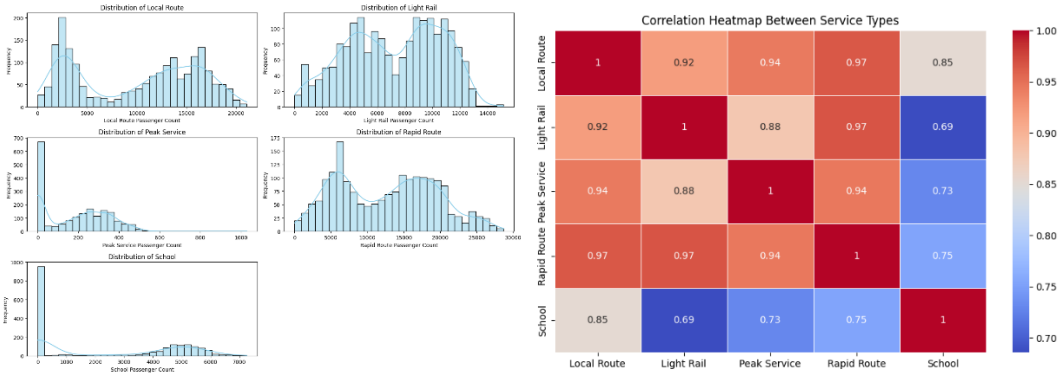
# Forecasting Daily Public Transport Passenger Journeys

**Insights:**

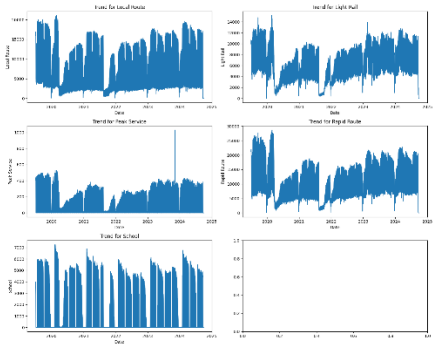



**Distribution plot**:

- **Local & Rapid Routes:** Bimodal, peaks at 5,000 and 15,000 passengers, indicating varied demand.
- **Light Rail:** Normal distribution, centered around 7,000–8,000 passengers, stable usage.
- **Peak Service & School:** Highly skewed, most counts <1,000, serving niche needs.
- **Overall**: Varied demand across services; Light Rail is the most consistent.

**Correlation Heatmap:**

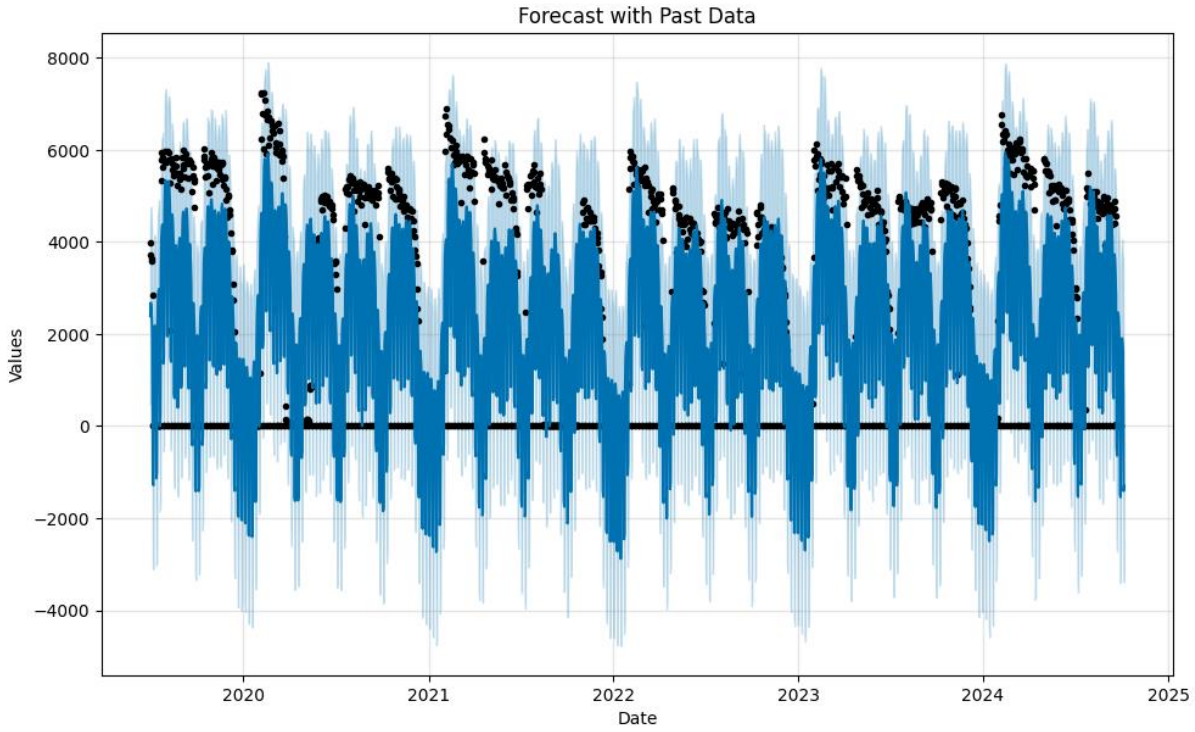
- Local, Rapid, and Peak Services show strong correlations (0.94–0.97), indicating similar demand patterns. School services have weaker correlations (0.69–0.85), reflecting distinct usage trends.



**Line plot:**

- **Local Route and Light Rail:** Both exhibit significant fluctuations, with distinct drops, likely due to external factors like holidays or disruptions.
- **Rapid Route:** Similar patterns to "Local Route," with periodic peaks and troughs, indicating possible seasonal or demand-based changes.
- **Peak Service:** Generally lower numbers with less variation, except for a notable spike in 2024.
- **School:** Regular periodicity, possibly corresponding to school sessions and breaks, with a clear drop-off during vacations.

**Forecasting with a past data:**




**Seasonality and Trend Analysis:**

- The historical data demonstrates strong seasonal patterns, with recurring peaks and troughs over time. This indicates cyclical behavior, likely influenced by weekly or seasonal factors such as weekdays, holidays, or specific operational schedules**.**

**Forecast Accuracy and Uncertainty:**

- The forecasted values closely follow the historical trend, showcasing the model's capability to capture periodic behavior effectively.
- The uncertainty bands (lighter blue areas) widen as the forecast extends further into the future, indicating decreasing confidence levels. This is typical in time-series forecasting and highlights the need for caution in long-term predictions.

**Critical Observations:**

- Peaks represent high-activity periods that could be associated with specific events or regular high-demand cycles.
- Troughs align with low-demand phases, potentially tied to weekends, holidays, or other interruptions.

This report describes using Prophet, an open-source forecasting tool developed by Facebook, for predicting daily public transport journeys. Prophet is ideal for time series data with seasonality and trend components, as it decomposes data into trend, seasonality, and holiday effects.

**Model parameters:**

**1.Growth Trend:**

- Type: Linear
- The model assumes a linear growth trend over time. (For datasets with saturation limits, logistic growth can be used.)

**2. Seasonality:**

- Yearly Seasonality
- Weekly Seasonality
- Daily Seasonality.

**3. Holidays:**

- Holidays: Not included. However, Prophet supports incorporating holiday effects for better accuracy when data shows holiday-related fluctuations.

**4. Changepoints:**

- Number of Changepoints
- Changepoint Prior Scale
- The model allows some flexibility for trend changes over time.

**5. Uncertainty Intervals:**

- Confidence Interval Width
- This means predictions have an 80% confidence level for their uncertainty bounds.

**6. Regularization:**

- Seasonality Prior Scale
- Holidays Prior Scale

**Results and Visualization:**

The forecasts were visualized to show trends, seasonal patterns, and holiday deviations, providing insights for each transport mode. Visualizations included confidence intervals and correlation heatmaps, highlighting demand fluctuations.