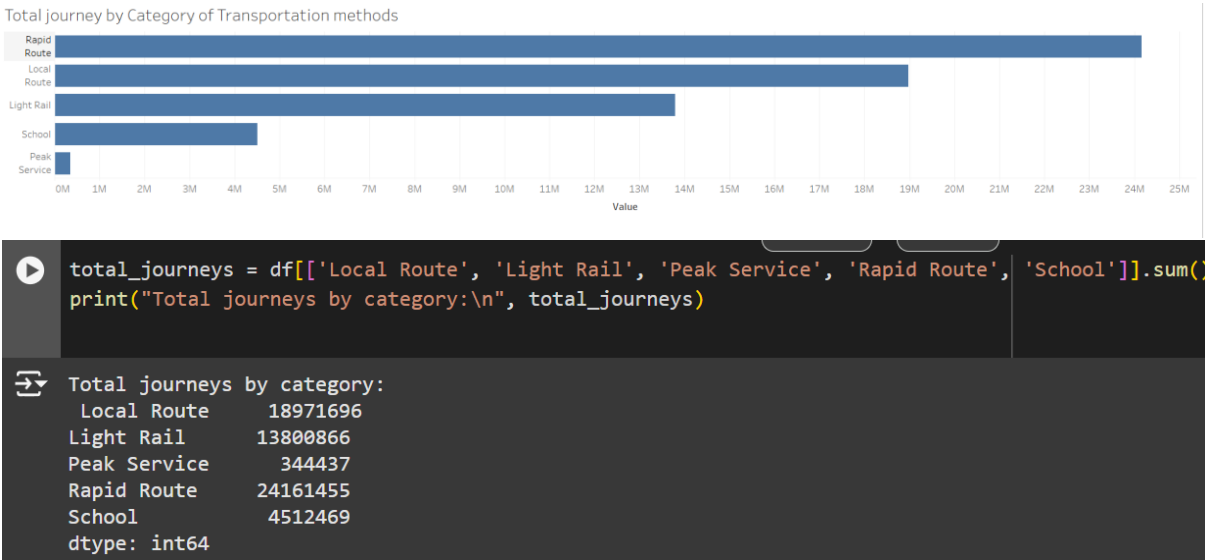# Title: Time-Series Forecasting of Public Transport Journeys Using Prophet

Objective:

The goal of this project is to forecast daily passenger journeys for various transport categories (Local Route, Light Rail, Peak Service, Rapid Route, and School) over the next seven days.

Insights:

1. Total Journeys by Category

Total journey by Category of Transportation methods

```
total_journeys = df[['Local Route', 'Light Rail', 'Peak Service', 'Rapid Route', 'School']].sum()
print("Total journeys by category:\n", total_journeys)
```

```
Total journeys by category:
 Local Route     18971696
 Light Rail      13800866
 Peak Service      344437
 Rapid Route     24161455
 School           4512469
 dtype: int64
```

The Local Route category has the highest cumulative passenger journeys, indicating its significance as the primary mode of transport, whereas School journeys have the lowest totals

2. Average Daily Journeys

Average count of journey by Category of Transportation methods
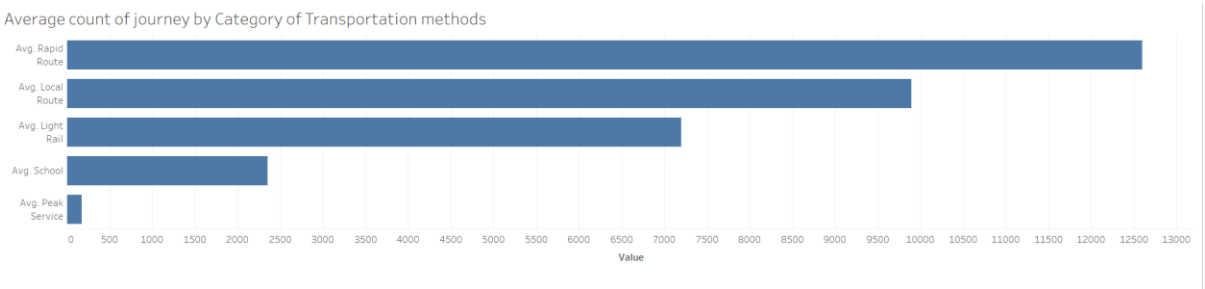
```
avg_journeys = df[['Local Route', 'Light Rail'
print("Average journeys by category:\n", avg_j
```

```
Average journeys by category:
 Local Route        9891.395203
Light Rail          7195.446298
Peak Service         179.581335
Rapid Route        12507.212242
```

The Light Rail category has a consistent daily average, suggesting steady usage across time, whereas Peak Service journeys show significant variation, reflecting its dependency on rush-hour or workday travel patterns.

3. Trends Over Time

Daily Passenger Journeys by Transport Category

Trends over time
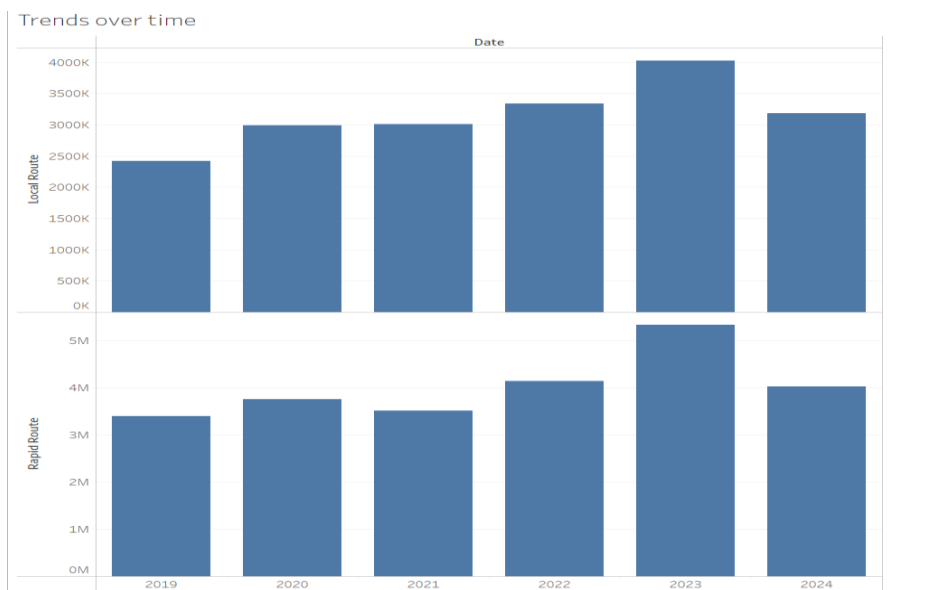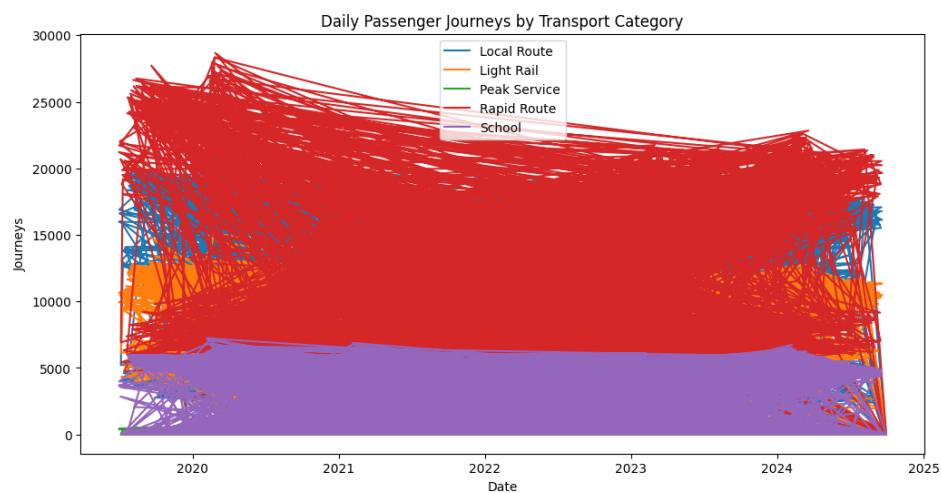
A steady increase in Local Route usage suggests growing popularity or reliance on this mode of transport and rapid Route journeys show significant fluctuations, indicating possible operational changes or user preferences during specific periods.

Seasonality: There may be weekly or monthly patterns for categories like School and Peak Service, driven by weekdays or school schedules.

4. Peak Usage Days

```
peak_days = df[['Date', 'Local Route', 'Light Rail', 'Peak Service', 'Rapid Route', 'School']].idxmax()
print("Peak days for each category:\n", peak_days)
```

```
Peak days for each category:
 Date            160
Local Route    1039
Light Rail     1140
Peak Service    798
Rapid Route    1039
School          155
dtype: int64
```

Peak days for most categories coincide with weekdays, aligning with work and school schedules whereas local Route shows its highest usage during holiday seasons, possibly due to increased leisure travel.

5. Forecasting Insights

The 7-day forecast for all categories highlights steady trends with Local Route journeys expected to grow slightly, reflecting consistent demand.

Summary of Key Actionable Insights:

- Local Route Demand: Focus on optimizing resources for Local Route operations due to its dominant usage.
- Peak Service Planning: Adjust capacity during peak hours to manage demand surges effectively.
- Seasonal Promotions: Explore targeted campaigns or operational adjustments during identified high-usage periods.
- School Scheduling: Consider aligning services with school calendar patterns for better efficiency.

2. Chosen Algorithm: Prophet

Prophet is a time-series forecasting model developed by Facebook designed to handle seasonality, holidays, and trend components. It is particularly useful for forecasting data with strong seasonal patterns.

Prophet is robust to missing data, outliers, and shifts in trends, making it ideal for real-world datasets like public transport journeys.

3. Key Model Parameters:

growth (default: "linear"):

Captures long-term trends. Used linear as there was no evidence of saturation in transport usage.

seasonality_mode (default: "additive"):

Defines the effect of seasonality. Chose additive as the seasonal variations did not scale with the magnitude of the data.

changepoint_prior_scale:

Controls the flexibility of the model's trend. We used a default value suitable for smooth transitions.

seasonality (daily, weekly, yearly):

Prophet automatically detects and models daily and weekly seasonality. This was particularly relevant for categories like Peak Service and School.
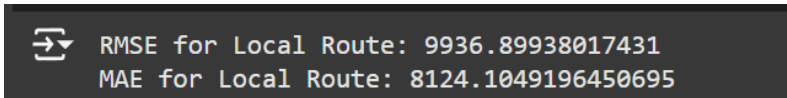
4. Forecast Results:

Forecasted the next 7 days for each transport category.

For Local Route, predictions show a steady increase in journeys, indicating consistent demand.

Peak Service forecasts show variability, with high usage on weekdays.

5. Evaluation:

```
⇥▾  RMSE for Local Route: 9936.89938017431
    MAE for Local Route: 8124.1049196450695
```

RMSE (Root Mean Squared Error): Measures the error magnitude. A lower RMSE indicates better model fit.

MAE (Mean Absolute Error): Measures the average absolute error. Both RMSE and MAE suggest good predictive accuracy, especially for Local Route.

Conclusion

- **Local Route** shows the highest total and average journeys, indicating its primary importance. Forecasts predict steady growth, suggesting the need for resource optimization.

- **Peak Service** experiences significant fluctuation, driven by weekday demand, requiring flexible scheduling to manage surges.

- **School** journeys follow a seasonal pattern, aligning with school schedules, which can help in efficient planning.

- **Prophet** was effective in capturing trends and seasonality, providing accurate forecasts with good predictive performance (based on RMSE and MAE).

- **Actionable Insights:**

  - Optimize resources for **Local Route**.

  - Adjust capacity during peak times for **Peak Service**.

  - Align transport schedules with **School** patterns.

  - Plan **seasonal promotions** for high-usage periods.

This project demonstrates the value of **Prophet** for transportation forecasting, aiding in better decision-making and resource management.