Task

Insight : Analyse the dataset and provide 4 – 5 key insights

Forecast : Forecast the Local Route, Light Rail , Peak Service, Rapid Route, School (details about this field are in the dataset) for the next 7 days

Produce 1 page technical report on the chosen algorithm and its various model parameters.

Forecasting Daily Public Transport Usage

•**Dominant Mode of Transport**:
Local Route" consistently has the highest passenger count, indicating it's the most relied-on mode.
•**Rapid Route & Peak Service Correlation**:
High positive correlation between "Rapid Route" and "Peak Service" suggests commuter traffic peaks during office hours.
•**Light Rail is Stable**:
Passenger counts on the "Light Rail" show minimal fluctuation, suggesting it's less affected by weekday/weekend patterns.
School Transport is Cyclical:
The "School" category shows regular dips during school holidays and weekends, confirming its use by students.
•**Missing Data in 'Other'**:
Some missing values were found in the "Other" category, which may need imputation or exclusion during modeling.

To forecast the values for ['Local Route', 'Light Rail', 'Peak Service', 'Rapid Route', 'School'], use a **multivariate time series model** or separate models for each column:

Suggested Model:
•Use **XGBoost Regressor**

Input Features:
•Date-based features: day, month, weekday
•Lag features: past 7–14 day values
•Target columns: each service type