

# Daily Public Transport Passenger Journeys by Service Type

NO OF ROWS

1,918

NO OF COLUMNS

7

The dataset titled "**Daily Public Transport Passenger Journeys by Service Type**" contains data about the number of passenger journeys made on different types of public transport services on various dates.

**Data covers:** From **JULY 1, 2019** to **SEPTEMBER 29, 2024** -USING POWERBI SORTING.

## 1.COVID-19 Impact

- Notable dips in passenger journeys during:
  - **March–May 2020**
- Demonstrates how external events disrupt usage and how recovery trends can be measured.

## 2. School Transport Seasonality

- School journeys are highly seasonal, almost disappearing in school breaks and holidays.

## 3. Local Route Dominance

- **Local Routes consistently contribute the highest share** of total passengers across all service types.
- Often accounts for **over 50%** of the total daily journeys — a critical backbone of public transport.

## 4. Total Daily Load Stability

Despite fluctuations in individual services, the overall total passenger count remains surprisingly stable on weekdays.

## 5. Weekend Dips

All service types show **notable drops in ridership on weekends**, presenting an opportunity to **adjust fleet operations** and reduce idle capacity

STATISTICAL INFORMATION

Detected missing values and outliers-using IQR.

#### Missing & Outliers Summary:

	Missing Values	Outliers
Date	0	0
Local Route	0	0
Light Rail	0	0
Peak Service	0	1
Rapid Route	0	0
School	0	0
Other	20	9

#### Algorithm to be used

The algorithm i have chosen is **SARIMA**

### What is SARIMA?

SARIMA stands for Seasonal AutoRegressive Integrated Moving Average. It is a time series forecasting model that extends ARIMA to handle seasonality in data.

### Why use SARIMA for this dataset?

The dataset contains **daily passenger counts** for different public transport service types. This kind of data usually exhibits:

- **Trend** (e.g., increasing/decreasing usage over time),
- **Seasonality** (e.g., more usage on weekdays vs. weekends),
- **Autocorrelation** (today's value is related to yesterday's).

SARIMA is ideal because:

- It handles **both trend and seasonality**
- Captures **weekly patterns** (we set  $m=7$  for weekly seasonality),
- Produces **accurate short-term forecasts**, such as 14-day predictions.

