Metropolitan Transportation Authority (MTA) Ridership Analysis Post-Pandemic:

Project Documentation

Table of Contents

---

1. Introduction

The COVID-19 pandemic has significantly impacted public transportation systems worldwide. The Metropolitan Transportation Authority (MTA) in New York City saw a drastic drop in ridership during the peak of the pandemic. As the city began to recover, understanding how ridership patterns evolved became crucial for decision-makers to optimize services, allocate resources, and plan for future needs.

This project aims to analyze the **MTA daily ridership data** to track recovery trends across different transportation modes, including subways, buses, Long Island Rail Road (LIRR), Metro-North, and bridges/tunnels.

---

2. Project Objectives

The main objectives of this project are to:

- Identify the highest ridership days for each mode of transportation.
- Analyze ridership recovery trends compared to pre-pandemic levels.
- Calculate averages and totals for different transportation modes to understand long-term trends.
- Provide actionable insights into how various transportation services have rebounded.
- Support the MTA in strategic planning and resource allocation by identifying key recovery patterns.

---

3. Database Structure

Database and Table Setup

We created a database named **`mtaRaider`** to store and analyze ridership data. The database includes a table named `raider` with the following fields:

- **Date_MTA**: Date of the data entry.
- **Subways_Total_Estimated_Ridership**: Total subway ridership on the given date.
- **Subways_Percentage_of_Pre_Pandemic**: Subway ridership as a percentage of pre-pandemic levels.
- **Buses_Total_Estimated_Ridership**: Total bus ridership.
- **Buses_Percentage_of_Pre_Pandemic**: Bus ridership as a percentage of pre-pandemic levels.
- **LIRR_Total_Estimated_Ridership**: Total ridership for Long Island Rail Road.
- **LIRR_Percentage_of_Pre_Pandemic**: LIRR ridership as a percentage of pre-pandemic levels.
- **MetroNorth_Total_Estimated_Ridership**: Total ridership for Metro-North.
- **MetroNorth_Percentage_of_Pre_Pandemic**: Metro-North ridership as a percentage of pre-pandemic levels.
- **AccessARide_Total_Scheduled_Trips**: Total scheduled trips for Access-A-Ride.
- **AccessARide_Percentage_of_Pre_Pandemic**: Access-A-Ride ridership as a percentage of pre-pandemic levels.
- **Bridges_Tunnels_Total_Traffic**: Total traffic through bridges and tunnels.
- **Bridges_Tunnels_Percentage_of_Pre_Pandemic**: Traffic percentage compared to pre-pandemic levels.
- **Staten_Island_Railway_Total_Estimated_Ridership**: Ridership for Staten Island Railway.
- **Staten_Island_Railway_Percentage_of_Pre_Pandemic**: Percentage recovery for Staten Island Railway.

---

4. Data Analysis

4.1 Highest Ridership Days

We identified the day with the highest ridership for subways using the following query:

```
SELECT Date_MTA, Subways_Total_Estimated_Ridership
FROM mta_daily_ridership
ORDER BY Subways_Total_Estimated_Ridership DESC
LIMIT 1;
```

- **Result**: The highest ridership day since the pandemic was recorded on [specific date] with [X] passengers.

4.2 Average Bus Ridership Recovery

To understand the bus ridership recovery, we calculated the average percentage compared to pre-pandemic levels:

```
SELECT AVG(Buses_Percentage_of_Pre_Pandemic) AS Avg_Bus_Per
FROM mta_daily_ridership;
```

- **Insight**: The average bus ridership was approximately **Y%** of pre-pandemic levels, indicating gradual recovery.

4.3 Metro-North Ridership Analysis

Identifying days when **Metro-North** ridership exceeded 100,000 helps in understanding peak demand periods:

```
SELECT Date_MTA, MetroNorth_Total_Estimated_Ridership
FROM mta_daily_ridership
WHERE MetroNorth_Total_Estimated_Ridership > 100000;
```

- **Finding**: The Metro-North system saw significant ridership on [list of dates].

4.4 Weekly Ridership Trends

We calculated weekly averages to smooth out daily fluctuations and identify broader patterns:

```
SELECT YEAR(Date_MTA) AS Year, WEEK(Date_MTA) AS Week,
       AVG(Subways_Total_Estimated_Ridership) AS Avg_Subway_Ridership
FROM mta_daily_ridership
GROUP BY Year, Week;
```

- **Result**: Weekly trends show that ridership is gradually approaching pre-pandemic levels in [specific months].

4.5 Railway Ridership Recovery Comparison

We monitored Staten Island Railway's recovery by comparing day-to-day changes:

```
SELECT Date_MTA,
       (Staten_Island_Railway_Percentage_of_Pre_Pandemic -
        LAG(Staten_Island_Railway_Percentage_of_Pre_Pandemic)
        OVER (ORDER BY Date_MTA)) AS Daily_Change
FROM mta_daily_ridership;
```

- **Insight**: Ridership recovery was consistent, with occasional spikes indicating special events or holidays.

4.6 **Yearly Total Ridership Across Multiple Modes**

```
sql
SELECT YEAR(Date_MTA) AS Year,
       SUM(Subways_Total_Estimated_Ridership) AS Total_Subway_Ridership,
       SUM(Buses_Total_Estimated_Ridership) AS Total_Bus_Ridership,
       SUM(LIRR_Total_Estimated_Ridership) AS Total_LIRR_Ridership,
       SUM(MetroNorth_Total_Estimated_Ridership) AS Total_MetroNorth_Ridership
FROM mta_daily_ridership
GROUP BY YEAR(Date_MTA)
ORDER BY Year;
```

- Summarizes **yearly ridership** for subways, buses, LIRR, and Metro-North to see how each mode has recovered annually.

4.7 **Top 3 Highest Subway Ridership Days per Year**

```
SELECT Date_MTA,
      YEAR(Date_MTA) AS Year,
      Subways_Total_Estimated_Ridership
   FROM
       mta_daily_ridership AS outer_table
   WHERE
       (SELECT COUNT(DISTINCT Subways_Total_Estimated_Ridership)
        FROM mta_daily_ridership AS inner_table
        WHERE YEAR(inner_table.Date_MTA) = YEAR(outer_table.Date_MTA)
          AND inner_table.Subways_Total_Estimated_Ridership >=
   outer_table.Subways_Total_Estimated_Ridership) <= 3
   ORDER BY
   Year DESC , Subways_Total_Estimated_Ridership ASC;
```

Extracts the **top 3 ridership days** for each year to highlight peak ridership days.

---

5. Key Insights and Findings

- **Subways and Buses**: Subways have shown a faster recovery compared to buses, with ridership levels approaching nearly **Z%** of pre-pandemic levels.
- **Bridges and Tunnels**: Traffic through bridges and tunnels has almost fully recovered, indicating a shift toward private vehicle usage.
- **Access-A-Ride**: Recovery for Access-A-Ride services was slower, indicating that vulnerable populations may still face mobility challenges.

---

6. Conclusion

The analysis of MTA ridership data provides valuable insights into post-pandemic recovery patterns across different transportation modes. Understanding these trends can help policymakers and transportation authorities optimize services and plan for future needs. Subways and railroads, being critical to New York City's public transportation network, have shown a resilient recovery, while buses and specialized services like Access-A-Ride continue to face challenges.

---

7. Future Enhancements

- **Data Enrichment**: Integrate weather data, events, or COVID-19 case rates to understand external factors affecting ridership.
- **Predictive Analysis**: Use machine learning to forecast future ridership trends based on historical data.
- **Dashboard Visualization**: Develop interactive dashboards using tools like Power BI or Tableau for real-time monitoring of ridership data.