Problem Definition and Research Questions

Project Title: Analyzing and Forecasting Public Transportation Ridership Trends in New York (MTA)

Team Members:

- Gaser Ahmed Mohammed Saad
- Mohamed Abdelfattah Ibrahim Elwa
- Abdulrahman Mohamed Safaa Ismail
- Basant Abdelhalim Mohamed Safan
- Doha Medhat Elsayed Salem

Supervisor: Sherihan Ali

Date: September 19, 2025

1. Introduction

In the aftermath of the COVID-19 pandemic, public transportation systems worldwide faced unprecedented challenges with dramatic ridership declines. The Metropolitan Transportation Authority (MTA) of New York, being the largest transportation network in North America, experienced significant impacts across all its services. This project aims to analyze daily ridership patterns across MTA's comprehensive transportation network to understand recovery trends, identify operational insights, and support strategic decision-making for future transportation planning.

2. Business Problem Statement

The MTA is facing critical challenges in understanding and optimizing post-pandemic ridership recovery across its diverse transportation modes. Key issues include uneven recovery rates across different transportation services, fluctuating daily ridership patterns, potential operational inefficiencies, and the need to adapt services to new passenger behavior patterns. Understanding these dynamics is essential for resource allocation, service optimization, and strategic planning to ensure sustainable public transportation operations.

3. Research Questions

To address the identified transportation challenges, the project focuses on answering the following key research questions:

- How has ridership recovery varied across different MTA transportation modes (Subway, Bus, LIRR, Metro-North, etc.)?
- What are the seasonal and weekly patterns in ridership across all MTA services?
- Which transportation modes have shown the strongest recovery compared to pre-pandemic levels?
- How do ridership patterns correlate between different transportation services?

- What factors contribute to ridership fluctuations in the post-pandemic era?
- Can we predict future ridership trends based on historical recovery patterns?

4. Objectives

- Analyze ridership recovery patterns across all MTA transportation modes
- Identify seasonal, weekly, and daily trends in passenger behavior
- Compare pre-pandemic vs. current ridership performance across services
- Develop insights for optimizing transportation service allocation
- Create predictive models for future ridership estimation
- Provide data-driven recommendations for strategic transportation planning

5. Scope & Limitations

Scope:

- The analysis will be based on MTA daily ridership data from March 2020 onwards
- Covers all major MTA services: Subways, Buses, LIRR, Metro-North, Access-A-Ride, Bridges & Tunnels, and Staten Island Railway
- Focus on descriptive, diagnostic, and predictive analytics

Limitations:

- Analysis limited to available MTA ridership data without external economic indicators
- Some data gaps may exist for certain services during specific periods

- External factors (weather, major events, policy changes) are not directly included in the dataset
- The study focuses on systemwide data rather than station-specific or route-specific analysis

6. Key Performance Indicators (KPIs)

To measure the success of the analysis and track transportation recovery, the following KPIs will be used:

#	KPI	Purpose	Formula / DAX Expression
1	Total Daily Ridership	Measure the total number of riders across all modes per day	Total_Ridership = SUM(Subway) + SUM(Bus) + SUM(LIRR) + SUM(MetroNorth) + SUM(SIR) + SUM(AAR)
2	Recovery Rate (% vs Pre-Pandemic)	Compare current ridership with pre-pandemic baseline	Recovery% = DIVIDE([Total_Ridership], [PreCOVID_Ridership]) * 100
3	Top Transport Mode	Identify the most used transport mode on a given day/week/month	TOPN(1, SUM(Ridership), Mode)
4	Monthly Growth Rate	Track month-over-month change in ridership	(ThisMonth - LastMonth) / LastMonth
5	Ridership Share by Mode (%)	Distribution of ridership across modes	SUM(Ridership by Mode) / [Total_Ridership] * 100
6	Daily Variance vs Monthly Average	Detect anomalies compared to the monthly average	Ridership_Today - AVERAGEX(VALUES(Month), [Total_Ridership])
7	Bridge & Tunnel Utilization (%)	Evaluate traffic recovery levels vs pre-pandemic	DIVIDE([Traffic], [PreCOVID_Traffic]) * 100
8	Week-over-Week Change (%)	Compare current week ridership with previous week	(ThisWeek - LastWeek) / LastWeek
9	Forecasted Ridership (Next Month)	Predict next month's ridership using time series forecasting	Power BI Forecast Model
10	Mode Dependency Index	Show dependency on a single mode (e.g., Subway)	Ridership_Mode / [Total_Ridership]
11	Peak Ridership Day	Identify the highest ridership day within a time range	MAX([Total_Ridership])

12	Mode-Specific Recovery Rate	Recovery % for each transport mode individually	SUM(Ridership_Mode) / SUM(PreCOVID_Mode)
13	Access-A-Ride Utilization (%)	Measure recovery for AAR (paratransit) users	SUM(AAR) / SUM(AAR_PreCOVID)
14	Year-over-Year Growth Rate (YoY)	Compare ridership with the same period last year	(ThisYear - LastYear) / LastYear
15	Weekend vs Weekday Ratio	Measure how demand differs between weekdays and weekends	Weekend_Ridership / Weekday_Ridership

7. Expected Outcomes

- **Comprehensive Recovery Analysis:** Clear insights into which transportation modes have recovered most effectively and which need additional focus
- **Ridership Pattern Identification:** Understanding of daily, weekly, and seasonal patterns that can inform service scheduling and resource allocation
- **Predictive Capabilities:** Reliable forecasting models for future ridership estimation to support strategic planning
- **Operational Recommendations:** Data-driven suggestions for optimizing service frequency, capacity allocation, and route planning
- **Strategic Planning Support:** Evidence-based recommendations for long-term transportation infrastructure and service development
- **Performance Benchmarking:** Established baselines and targets for measuring future transportation performance and recovery progress