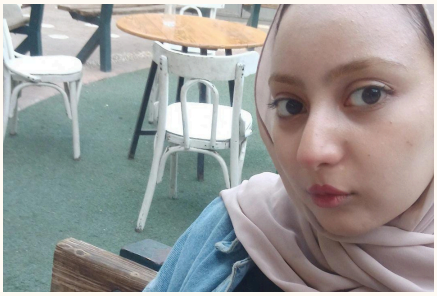
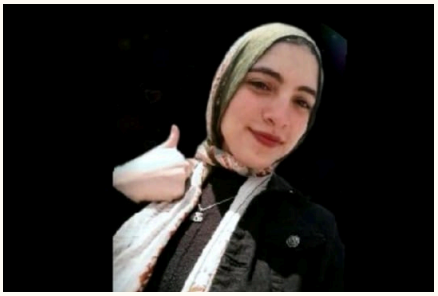
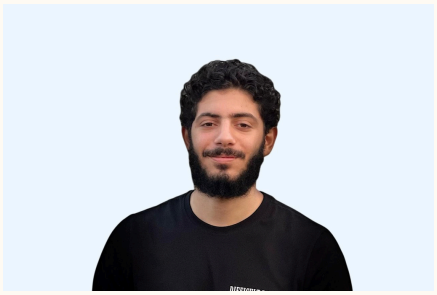
MTA Daily Ridership

Our Team

		
Sarah Emad	Tasneem Mohsen	Alyaa Ibrahim

Our Team

		
Fatma Hassan	Momen Haithem	Galal Eldeen Mohamed

Overview

This project aims to analyze **MTA daily ridership data** to identify trends in subway, bus, and commuter rail usage. The goal is to assess ridership recovery post-pandemic, understand peak travel times, and explore factors affecting transit usage. The dataset includes daily ridership counts for subways, buses, railroads, and bridges/tunnels, along with percentage comparisons to pre-pandemic levels. The dataset includes daily ridership counts from 2020 onward for subways, buses, railroads, and bridges/tunnels. It is sourced from MTA Open Data, consisting of millions of data points aggregated from MetroCard/OMNY tap-ins and automated passenger counters.

Objectives

- Identify peak ridership hours for different transportation modes.
- Analyze subway, bus, and commuter rail ridership trends over time.
- Compare current ridership to pre-pandemic levels to assess recovery.
- Evaluate the impact of external factors (e.g., COVID-19, weather, economy) on ridership.
- Provide actionable insights for **transportation planning and policy-making**.

Scope

- Data collection and cleaning from MTA ridership records.
- Exploratory data analysis (EDA) and visualization.
- Identifying key patterns and trends in ridership data.
- Forecasting ridership trends using machine learning models.
- Developing an interactive dashboard to visualize insights.

Data description

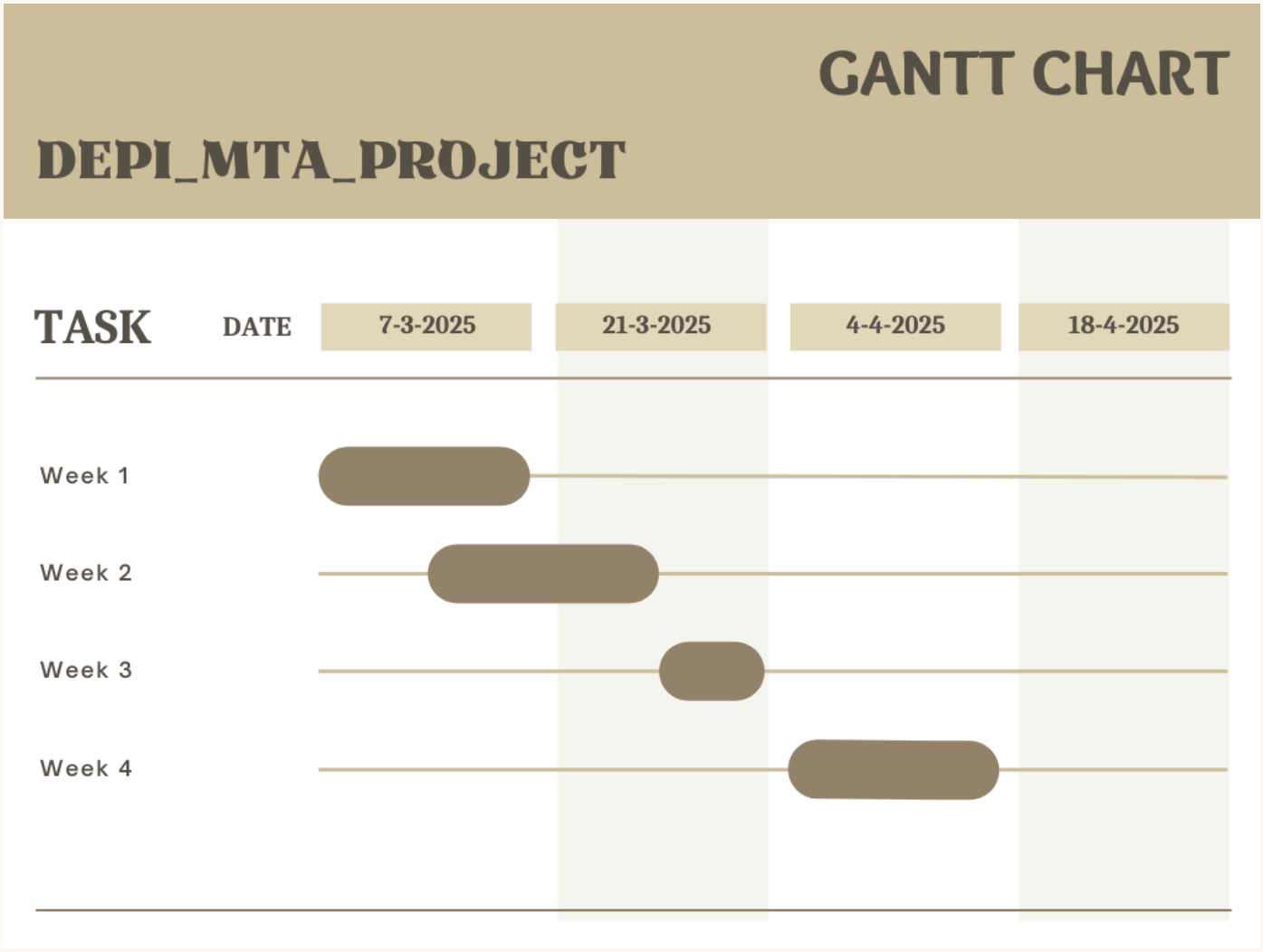
- Dataset Name: MTA Daily Ridership
- Source: Metropolitan Transportation Authority (MTA)
- Date Range: Starting from March 1, 2020
- Granularity: Daily
- Fields:

1. Date
2. Total Estimated Ridership for:
    - Subways
    - Buses
    - LIRR (Long Island Rail Road)
    - Metro-North Railroad
    - Access-A-Ride (paratransit)
    - Bridges and Tunnels traffic
    - Staten Island Railway
3. Corresponding % of Pre-Pandemic Ridership

Data Modeling & System Design

- **Technologies:** Python (Pandas, Matplotlib, Seaborn, Plotly), SQL, Power BI, Tableau.
- **Data Processing:** Cleaning, handling missing values, and transformation.
- **Visualization Tools:** Power BI, Tableau, and Python.
- **Machine Learning Models:** Time Series Forecasting (Prophet, ARIMA, or XGBoost).

Data Modeling



Stakeholder Analysis

Stakeholder	Role	Interest
MTA Operations Team	Service planners	Optimize transit scheduling and reduce costs
City Government	Policy makers	Ensure effective use of public funds
Commuters	End users	Reliable and safe transit options
Data Analysts/Researchers	Insight developers	Provide data-driven recommendations
Budget Office	Finance managers	Align transit funding with ridership demand

Key Performance Indicators (KPIs)

Metric	Success Indicator
Peak Ridership Hours Identified	Determine the busiest travel hours
Subway vs. Bus vs. Commuter Rail Trends	Compare ridership across transit modes
Pre vs. Post-Pandemic Recovery Rate	Assess recovery percentage since COVID-19
Forecast Accuracy for Ridership Trends	Measure prediction accuracy for future trends
Interactive Data Visualization Completed	Ensure clear, insightful, and interactive graphs

Insights from our Analysis

1- Post-COVID Recovery

- Some transportation modes (e.g., subway or bus) likely recovered faster than others like LIRR or Metro-North.
- Normalized plots probably show Subway rebounding sooner due to higher dependence in urban areas.

2- Weekend vs Weekday Trends

- Ridership patterns may differ significantly between weekdays and weekends.
- Recommendations might include adjusting service frequency or staffing based on observed usage.
- Weekend ridership is lower across all modes.
- Transit demand is highest on weekdays, suggesting potential for off-peak scheduling savings.

3- Seasonal or Temporal Trends

- Monthly or yearly breakdowns could reveal seasonal drops (e.g., holidays) or long-term recovery trends

4- Total Ridership Trends

- Identification of overall rising or falling ridership trends post-2020.
- Important for capacity planning and infrastructure investments.

5- Forecasting

- If ARIMA was applied to total_ridership, the model may project future demand, helping in planning service levels.

6- Mode Recovery Trends

From normalized ridership trends:

- Subways and buses likely showed faster recovery post-COVID.
- Commuter rails (LIRR, Metro-North) lagged—likely due to ongoing hybrid/remote work reducing daily suburban commutes.

7- Total Ridership Rebound

- Total ridership shows a gradual upward trend, which is a strong signal of recovery.
    - Fluctuations may align with seasonal patterns or COVID variant surges.

Recommendations You Could Derive

1- Resource Allocation

Increase service where recovery is strongest (e.g., subways on weekdays), reduce frequency for underused services.

2- Targeted Marketing

Promote LIRR or Metro-North through pricing or service incentives to stimulate slower-recovering ridership.

3- Policy Interventions

Consider flexible schedules or hybrid work policies to distribute peak demand more evenly.

4- Forecast-Driven Planning

Use ARIMA model forecasts for staffing, maintenance, and budgeting over the coming quarters.

> “Post-Pandemic Ridership Trends and Forecasting for MTA Transit Modes.”