**Project Name: MTA Daily Ridership Analysis**

**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.

**📅 Project Plan**

**Milestones & Timeline**

|  |  |
| --- | --- |
| **Milestone** | **Expected Completion** |
| Data Cleaning & Preprocessing | Completed |
| Exploratory Data Analysis (EDA) | Completed |
| Data Visualization & Trend Analysis | Ongoing |
| Forecasting Phase | 14-Apr-2025 |
| Interactive Dashboard Development | 21-Apr-2025 |
| Final Presentation & Report | 1-May-2025 |

**Deliverables**

* **Cleaned & Processed Dataset** → Ready for analysis.
* **Exploratory Data Analysis Report** → Summary of key trends.
* **Visualizations & Dashboards** → Interactive graphs to illustrate findings.
* **Forecasting Model** → Predictions for future ridership trends.
* **Final Report & Presentation** → Summary of insights & recommendations.

**Timeline**

A graph on a chart

AI-generated content may be incorrect.

**📝 Task Assignment & Roles**

|  |  |
| --- | --- |
| **Task** | **Assigned To** |
| Data Cleaning & Preprocessing Using SQL | Galal Eldeen Mohamed, Momen Haitham |
| Data Cleaning & Preprocessing Using Python | Sarah Emad |
| Data Analysis Using Python | Tasneem Mohsen |
| Forecasting using Machine Learning | Sarah Emad, Tasneem Mohsen |
| Visualization Dashboard using Power BI | Fatma |
| Visualization Dashboard using Tableau | Alyaa Ibrahim |

**📊 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. |

**🔍 Stakeholder Analysis**

|  |  |  |
| --- | --- | --- |
| **Stakeholder** | **Role in Project** | **Importance** |
| MTA Operators | Improve scheduling & capacity planning | High |
| Passengers | Understand transit trends & recovery | High |
| Urban Planners | Optimize public transport systems | Medium |
| Financial Analysts | Assess revenue impact of ridership changes | Medium |
| Government Agencies | Policy-making for public transit | Medium |

**Data Modeling & System Design**

* **Data Sources:** MTA Daily Ridership dataset (CSV format).
* **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).

**Final Deliverables by 1-May-2025**

✅ **Cleaned Data & Exploratory Analysis Report**  
✅ **Interactive Dashboard (Power BI / Tableau)**  
✅ **Predictive Model for Ridership Forecasting**  
✅ **Final Report & Presentation**

**Data Modeling**

A screenshot of a computer

AI-generated content may be incorrect.

This project will provide **data-driven insights** to improve MTA’s transit operations and aid in policy-making for a better public transportation experience. 🚆📊