

Final Project Proposal

# MTA Daily Ridership Analysis

Name	rule
<b>Kholoud Mohamed Swidan</b>	<ul style="list-style-type: none"> <li>-Team Leader</li> <li>-Defined key business questions and answered them using SQL</li> <li>- Designed and analyzed the first dashboard view</li> <li>- Helped shape the dashboard storytelling flow</li> </ul>
<b>Roaa Ahmed El Temamy</b>	<ul style="list-style-type: none"> <li>-Cleaned and prepared the dataset using Python</li> <li>-Designed and analyzed the second dashboard view</li> <li>-Helped shape the dashboard storytelling flow</li> </ul>
<b>Mariam Mohamed Gad</b>	<ul style="list-style-type: none"> <li>-Formulated business questions and answered them using Python</li> <li>-Created dashboard sketches</li> <li>-Designed the interface using Figma</li> <li>-Designed the third dashboard view</li> </ul>
<b>Rana Mohamed AbuElmaaty</b>	<ul style="list-style-type: none"> <li>- Performed data modeling for the project</li> <li>-Conducted the analysis and interpretation for the third dashboard view</li> </ul>
<b>Alaa Mohamed Abdelmonem</b>	

- **Problem:**

The COVID-19 pandemic significantly impacted public transportation usage, resulting in dramatic shifts in ridership trends and behavior. There's a need to understand these changes to support effective transit planning and recovery.

- **Proposed Solution:**

This project analyzes the MTA daily ridership dataset to identify trends before, during, and after COVID-19. It provides a visual and analytical understanding of ridership patterns using dashboards, forecasting, and interactive visualizations.

- **Unique Value Proposition:**

The project combines data analysis, predictive modeling, and intuitive dashboards to help stakeholders monitor performance, evaluate recovery, and optimize transportation policies. Objectives

- Objectives

1. Clean and preprocess the dataset to ensure accuracy and consistency.
2. Identify key analytical questions relevant to business needs.
3. Create an interactive dashboard for data visualization and reporting.
4. Deliver a comprehensive final report with actionable recommendations.

- Tools & Technologies

- SQL (SQL Server DBMS)
- Python (Pandas, Matplotlib, Scikit-learn)
- Data Modeling (Model views + ERD)
- Tableau Desktop
- Figma for Design

- **Milestones & Deadlines**

- **Week 1: Build Data Model, Data Cleaning and Preprocessing**

- **Tasks:**

- Data Preprocessing: Build a data model and clean and preprocess the data.
- Tools: SQL, Python (pandas, Matplotlib).

- **Deliverables:**

- Cleaned dataset ready for analysis.
- Data preprocessing notebook.
- Data Model:
- The original dataset was split into three separate files to facilitate better data organization and modeling:

Transportation Types and Modes File:

This file contains information about the different types and modes of transportation available in the dataset. It serves as a dimension table that helps in categorizing and analyzing the ridership data based on transport methods.

Total Daily Ridership File:

This file includes the total number of riders for each day. It acts as the main fact table, capturing the daily usage of the transportation system across all modes.

#### Date Dimension File:

This file provides a structured calendar table where each date is broken down into multiple components such as week number, month, year, and day of the week. This table supports time-based analysis and enables filtering and aggregation by various date attributes.

### - Week 2: Analysis Questions Phase

#### ● Tasks:

- Determine Data Analysis Questions: Determine all possible analysis questions that can be deducted from the given dataset and would be of interest to the organization's decision makers.

- Tools: SQL, Python (pandas, Matplotlib).

#### ● Deliverables:

- Set of analysis questions that can be answered via the dataset.

-Sample of Questions:

1-What is the total number of passengers across all transportation modes for each day?

2-What is the average daily ridership in the subway?

3-What are the top 5 busiest days in the subway?

4-How many days had subway ridership below 50% of pre-pandemic levels?

5-What is the percentage of subway and bus ridership compared to pre-pandemic levels for each day?

6-What are the top 5 busiest days across all transportation modes combined?

7-How does ridership change daily, weekly, and monthly?

8-Is there a seasonal pattern (e.g., summer vs. winter)?

9-What is the growth or decline rate of ridership over time?

10-How did the COVID-19 pandemic impact ridership?

11-Is there a difference between weekend and weekday ridership?

12-Which mode of transportation has the highest ridership?

13-Does an increase in car traffic reduce public transit ridership?

14-What are the busiest months for ridership?

15-How do different transit modes compare daily?

16-What is the Relationship Between Bridge Traffic and Public Transit Usage?

### **- Week 3: Dashboard Sketching Phase**

- Tasks:

- Initiate dashboard sketching by translating analytical goals into preliminary visual layouts.
- Define key metrics, chart types, and layout structure to ensure alignment with user needs and analytical objectives.
- Prepare a visual blueprint to guide the dashboard development process.

- Tools:

- Excalidraw

- Figma

- Deliverables:

- Initial dashboard sketches outlining visual structure and key components.
  - Documented rationale behind design decisions and selected visualizations.

- **Week 4: Visualization Dashboard and Final Presentation**

- **Tasks:**

- Build a Visualization Dashboard:

Develop an interactive Tableau dashboard that presents clear, insightful visualizations addressing all key analytical questions identified earlier. The dashboard was designed to include three comprehensive views, each tailored to highlight specific aspects of the dataset. Emphasis was placed on usability, visual clarity, and alignment with stakeholder needs, ensuring that the final product facilitates data-driven decision making.



- Prepare Final Presentation:

Create a structured final presentation and written report summarizing the full project workflow. This included data preprocessing, analysis methodologies, model structuring (if applicable), and dashboard design rationale. The deliverables were tailored for both technical and non-technical audiences, ensuring effective communication of insights and project impact.

- Tools:

- SQL, Python (pandas, Matplotlib) for data handling and exploratory analysis.
- Tableau for data visualization and dashboard development.
- PowerPoint or Google Slides for final presentation preparation.

- **Deliverables:**

- Visualization dashboard
- Final report and presentation

- **Key Performance Indicators (KPIs)**

- Accuracy of data preprocessing and cleaning.
- Number of valuable insights extracted from data analysis.
- Accuracy of forecasting models.
- User engagement and clarity of the visualization dashboard.
- Quality and structure of the final report and presentation.