Digital Egypt Pioneers Initiative – DEPI

مبادرة رواد مصر الرقمية

MTA DAILY RIDERSHIP

PROJECT DOCUMENTATION

ONL2_DAT2_G2

Group #3 Project

GROUP 3 MEMBERS

AHMED SHERIF

AHMED MOHAMED

AMER AHMED

EBTSAM EMAM

Introduction

Public transportation is a critical component of urban mobility, and understanding ridership trends can help transit authorities make data-driven decisions. The COVID-19 pandemic significantly impacted transit usage worldwide, including the Metropolitan Transportation Authority (MTA) network in New York. This project aims to analyze MTA daily ridership data, comparing post-pandemic recovery trends against pre-pandemic levels.

By examining subway, bus, commuter rail, and other transit services, we aim to uncover patterns that can inform infrastructure planning, service adjustments, and policy decisions.

Overview

This project analyzes the daily ridership of the Metropolitan Transportation Authority (MTA) across different transit modes, comparing post-pandemic ridership levels to pre-pandemic benchmarks. The goal is to identify trends, patterns, and insights that can inform transportation policies and operational decisions.

Tools and Resources

1. **Development Tools**
    - Power BI Desktop: Primary dashboard development platform
    - Canva:
        - Custom background design
        - Layout optimization
        - Performance enhancement through pre-designed elements
        - Brand-consistent visual elements
2. **AI Assistance**
    - *ChatGPT:*
        - Data analysis suggestions
        - DAX measure optimization
        - Documentation support
        - Problem-solving assistance
    - *Claude 3.5:*
        - Technical writing
        - Analysis validation
        - Feature recommendations
        - Documentation enhancement
3. **Design Resources**
    - *MTA Brand Guidelines:*
        - Official color palette
        - Logo usage
        - Typography standards
    - *Color Selection Tools:*
        - Color contrast checker (Eye-dropper)

4. **Data Processing**
    - Excel: Initial data cleaning and formatting
    - Power Query: Data transformation and modeling
    - DAX Studio: Measure optimization and testing
5. **Collaboration and Version Control**
    - GitHub: Code and documentation repository
    - Power BI Service: Dashboard publishing and sharing
    - Zoom and Google Teams Meetings:
    - Weekly team status updates:
        - Design review sessions
        - Technical problem-solving discussions
        - Knowledge sharing sessions

Objectives

This project aims to:

- **Analyze Ridership Trends**: Identify patterns in daily ridership across different transportation modes.
- **Compare Pre- and Post-Pandemic Ridership**: Evaluate how public transportation usage has recovered compared to pre-pandemic levels.
- **Discover Peak Usage Periods**: Determine the busiest times and days for different modes of transportation.
- **Identify Factors Influencing Ridership**: Explore potential correlations between ridership changes and external factors such as policy changes, seasonal variations, and economic conditions.
- **Provide Data-Driven Insights**: Generate meaningful visualizations and insights that can help stakeholders, policymakers, and transit planners make informed decisions.

Data Source and Description

The dataset used in this analysis is the **MTA Daily Ridership Data**, which provides daily ridership estimates for various MTA-operated transit services. The data includes subway, bus, Long Island Rail Road (LIRR), Metro-North, Access-A-Ride, and Bridges and Tunnels traffic counts. Additionally, the dataset includes percentage comparisons to equivalent pre-pandemic ridership levels.

Data Collection Methodology

- **Subway and Bus Ridership:** Derived from MetroCard and OMNY swipes and taps, with estimated cash fare additions.
- **LIRR and Metro-North Ridership:** Estimated using a model based on ticket sales.
- **Access-A-Ride Trips:** Reflects scheduled trips as an indicator of demand rather than completed trips.
- **Bridges and Tunnels Traffic:** Based on toll collection system data.
- **Pre-Pandemic Comparison:** Uses 2019 monthly averages for non-holiday weekdays, Saturdays, and Sundays/holidays.

Key Data Fields:

- **Date:** The date of travel (MM/DD/YYYY)
- **Total Estimated Ridership:** Daily estimated ridership for each transit mode
- **% of Comparable Pre-Pandemic Day:** The percentage of ridership compared to a pre-pandemic equivalent day

Ridership figures are determined using MetroCard and OMNY swipes, ticket sales, and automated passenger counters, with adjustments for estimated cash fares and scheduled trips. The dataset allows us to track ridership recovery trends over time, assess service demand, and identify patterns that impact transit usage.

Detailed columns description:

Data Label	Data Type	Data Description
Date	DATE	The date of travel (MM/DD/YYYY).
Subway Ridership	NUMERIC	The daily total estimated subway ridership.
Subway % of Pre-Pandemic	PERCENT	The daily ridership estimate as a percentage of subway ridership on an equivalent day prior to the COVID-19 pandemic.
Bus Ridership	NUMERIC	The daily total estimated bus ridership.
Bus % of Pre-Pandemic	PERCENT	The daily ridership estimate as a percentage of bus ridership on an equivalent day prior to the COVID-19 pandemic.
LIRR Ridership	NUMERIC	The daily total estimated LIRR ridership. Blank value indicates that the ridership data was not or is not currently available or applicable.
LIRR % of Pre-Pandemic	PERCENT	The daily ridership estimate as a percentage of LIRR ridership on an equivalent day prior to the COVID-19 pandemic.
Metro-North Ridership	NUMERIC	The daily total estimated Metro-North ridership. Blank value indicates that the ridership data was not or is not currently available or applicable.
Metro-North % of Pre-Pandemic	PERCENT	The daily ridership estimate as a percentage of Metro-North ridership on an equivalent day prior to the COVID-19 pandemic.
Access-A-Ride Trips	NUMERIC	The daily total scheduled Access-A-Ride trips. Blank value indicates that the ridership data was not or is not currently available or applicable.
Access-A-Ride % of Pre-Pandemic	PERCENT	The daily total scheduled trips as a percentage of total scheduled trips on an equivalent day prior to the COVID-19 pandemic. Blank value indicates that the ridership data was not or is not currently available or applicable.
Bridges and Tunnels Traffic	NUMERIC	The daily total Bridges and Tunnels traffic. Blank value indicates that the ridership data was not or is not currently available or applicable.
Bridges and Tunnels % of Pre-Pandemic	PERCENT	The daily total traffic as a percentage of total traffic on an equivalent day prior to the COVID-19 pandemic. Blank value indicates that the ridership data was not or is not currently available or applicable.
Staten Island Railway Ridership	NUMERIC	The daily total estimated SIR ridership.
Staten Island Railway % of Pre-Pandemic	PERCENT	The daily ridership estimate as a percentage of SIR ridership on an equivalent day prior to the COVID-19 pandemic.

Key Analytical Questions

Ridership Trends Over Time

1. How has the total daily ridership changed over time for each mode of transportation (subway, bus, LIRR, Metro-North, etc.)?
2. Are there noticeable seasonal patterns or trends in ridership?
3. How does ridership vary by day of the week?
4. Are there any anomalies or sudden changes in ridership, and can they be linked to external factors (e.g., policy changes, weather events, holidays)?

Post-Pandemic Recovery Analysis

5. How do current ridership levels compare to pre-pandemic levels across different transportation modes?
6. Which transportation modes have recovered the fastest, and which are still lagging behind?
7. Is there a difference in recovery rates between weekdays and weekends?

Comparison Across Modes of Transportation

8. Which transportation mode has the highest ridership recovery rate compared to pre-pandemic levels?
9. How does bus ridership compare to subway ridership in terms of recovery trends?
10. Is there a shift in preference between public transportation modes post-pandemic?

Impact of External Factors

11. How do extreme weather conditions or significant events impact daily ridership?
12. Did fare changes or policy adjustments (e.g., introduction of OMNY, congestion pricing) affect ridership patterns?
13. Is there a correlation between economic indicators (e.g., employment levels, inflation) and ridership trends?

Bridges and Tunnels Traffic Trends

14. How has the traffic volume on bridges and tunnels changed over time?
15. How does the traffic recovery rate compare to public transit recovery rates?

16. Are more people choosing personal vehicles over public transportation post-pandemic?

Accessibility and Service Usage

17. How has the usage of Access-A-Ride changed over time?
18. Are there specific trends in Access-A-Ride usage that indicate changes in accessibility needs?

Future Insights and Forecasting

19. Can we predict future ridership trends based on historical data?
20. What factors are most strongly correlated with ridership fluctuations, and can we model them for predictive analysis?

Technical Implementation and Data Modeling

Column Name Modifications

To enhance readability and usability, some column names have been shortened while preserving their original meaning. The table below outlines these modifications:

Original Column Name	Shortened Name
Date	Date
Subways: Total Estimated Ridership	Subway Ridership
Subways: % of Comparable Pre-Pandemic Day	Subway % of Pre-Pandemic
Buses: Total Estimated Ridership	Bus Ridership
Buses: % of Comparable Pre-Pandemic Day	Bus % of Pre-Pandemic
LIRR: Total Estimated Ridership	LIRR Ridership
LIRR: % of Comparable Pre-Pandemic Day	LIRR % of Pre-Pandemic
Metro-North: Total Estimated Ridership	Metro-North Ridership
Metro-North: % of Comparable Pre-Pandemic Day	Metro-North % of Pre-Pandemic
Access-A-Ride: Total Scheduled Trips	Access-A-Ride Trips
Access-A-Ride: % of Comparable Pre-Pandemic Day	Access-A-Ride % of Pre-Pandemic
Bridges and Tunnels: Total Traffic	Bridges and Tunnels Traffic
Bridges and Tunnels: % of Comparable Pre-Pandemic Day	Bridges and Tunnels % of Pre-Pandemic
Staten Island Railway: Total Traffic	Staten Island Railway Ridership
Staten Island Railway: % of Comparable Pre-Pandemic Day	Staten Island Railway % of Pre-Pandemic

These adjustments make the dataset easier to work with while maintaining clarity in analysis and visualization.

Data Structure Modifications

1. **Calculated Columns**
    - Total Ridership Column
        - Created as a calculated column to aggregate ridership across all transportation modes
        - Provides the foundation for percentage calculations and overall ridership analysis
        - Used in KPI calculations and trend analysis
2. **Date Dimension Table**
    - Created by duplicating the main table and extracting date-related information
    - *Added calculated columns:*
        - Year: Extracted from date for annual analysis
        - Month: Numeric month value
        - Month Name: Full month name for better readability
        - Day Name: Day of week for weekly pattern analysis
        - Quarter: Fiscal quarter designation
        - Season Name: Custom calculation for seasonal analysis
    - *This date table enables:*
        - Time intelligence calculations
        - Hierarchical time-based filtering
        - Seasonal and periodic trend analysis
3. **Percentage Analysis Table**
    - Created through table duplication and unpivoting
    - *Transformation process:*
        - Duplicated main data table
        - Unpivoted percentage columns
        - Created two resulting columns:

- Transport Types (categorical)
- Average Percentage (numerical)

- *Enables:*
  - Comparative analysis across transport modes
  - Simplified visualization of recovery trends
  - Streamlined percentage-based calculations

4. **Key Measures Implementation**
   Created DAX measures for modal analysis:

DAX Collapse

```
Bridges and Tunnels Percent =
DIVIDE(
    SUM('MTA_Daily_Ridership'[Bridges and Tunnels Traffic]),
    SUM('MTA_Daily_Ridership'[Total Ridership])
)

Bus Ridership Percent =
DIVIDE(
    SUM('MTA_Daily_Ridership'[Bus Ridership]),
     SUM('MTA_Daily_Ridership'[Total Ridership])
)

Subway Ridership Percent =
DIVIDE(
    SUM('MTA_Daily_Ridership'[Subway Ridership]),
    SUM('MTA_Daily_Ridership'[Total Ridership])
)

Access-A-Ride Percent =
DIVIDE(
    SUM('MTA_Daily_Ridership'[Access-A-Ride Trips]),
    SUM('MTA_Daily_Ridership'[Total Ridership])
)

LIRR Percent =
DIVIDE(
    SUM('MTA_Daily_Ridership'[LIRR Ridership]),
    SUM('MTA_Daily_Ridership'[Total Ridership])
)

Metro-North Percent =
DIVIDE(
    SUM('MTA_Daily_Ridership'[Metro-North Ridership]),
    SUM('MTA_Daily_Ridership'[Total Ridership])
)

Staten Island Railway Percent =
DIVIDE(
    SUM('MTA_Daily_Ridership'[Staten Island Railway Ridership]),
    SUM('MTA_Daily_Ridership'[Total Ridership])
)
```

These measures:

- Calculate the proportion of each transportation mode relative to total ridership

- Use SUM aggregation to handle multiple rows of data
- Implement DIVIDE function for safe division operations
- Support dynamic filtering and context transitions
- Drive key visualizations and KPIs throughout the dashboard

Data Relationships

- Main fact table (MTA_Daily_Ridership) connected to date dimension table
- Percentage analysis (Transport Types) table linked to main table
- Ensures consistent calculations across all visualizations
- Maintains data integrity and performance

These technical implementations enable:

- Accurate time-based analysis
- Reliable percentage calculations across different time periods
- Efficient data refreshes
- Improved dashboard performance
- Flexible filtering and analysis capabilities

Dashboard Design and Implementation

Design Elements and Branding

- The dashboard utilizes the official MTA brand colors, primarily featuring deep blues and light blue accents
- MTA logo is prominently displayed in the top-left corner of each page
- Custom-designed backgrounds created in Canva for optimal performance
- Consistent color scheme and layout across all pages for visual coherence
- Modern, clean interface with clear navigation elements

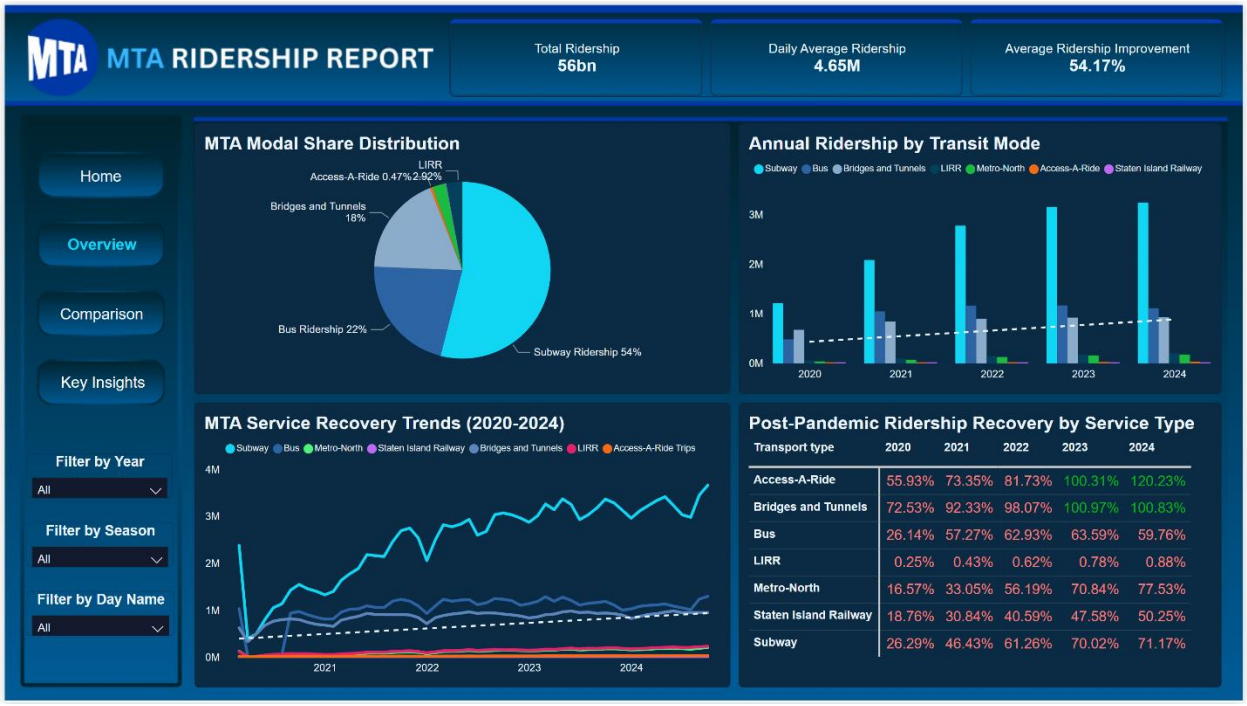
Navigation and Layout

The dashboard consists of four main pages:

1. **Home Page**

   - Provides comprehensive overview of MTA services
   - Contains detailed descriptions of each transit mode
   - Clean, informative layout with service definitions in bordered cards

2. Overview Page

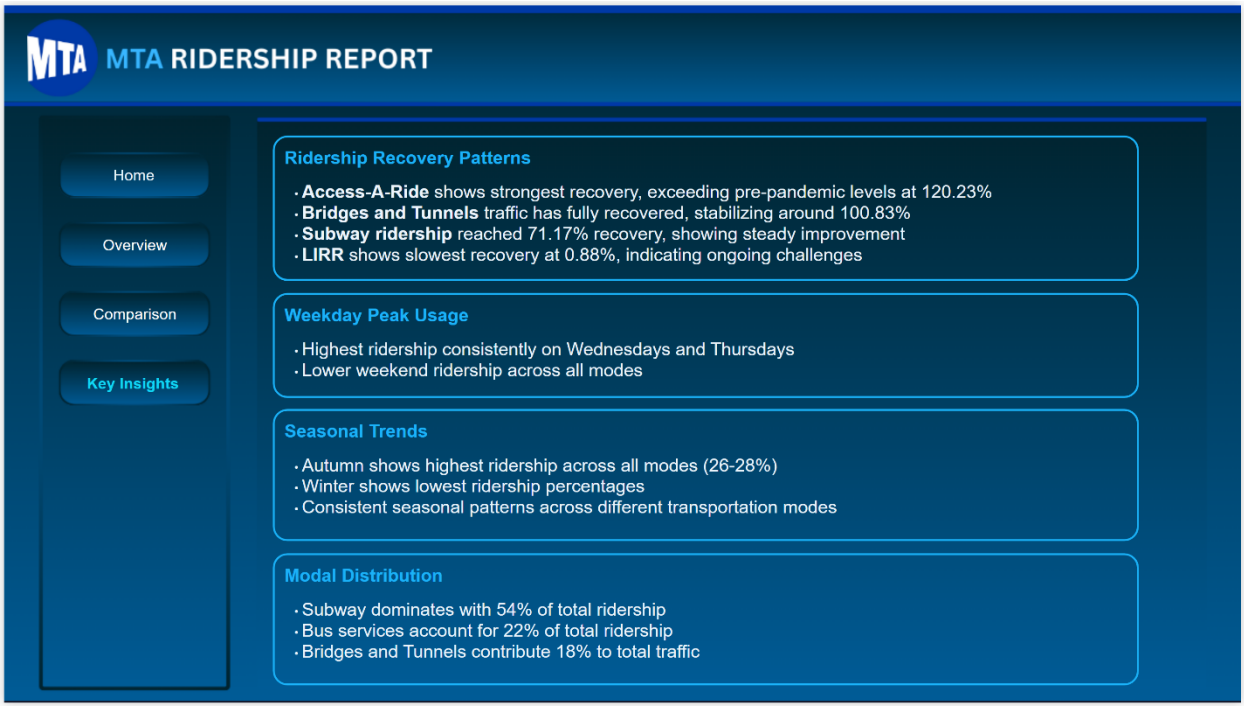


- *Features three main KPIs at the top:*
  - Total Ridership (56bn)
  - Daily Average Ridership (4.65M)
  - Average Ridership Improvement (54.17%)
- Year-over-year improvement table for all transport types
- Comparative ridership visualization across years
- Pie chart showing distribution of average ridership
- Line graph displaying post-pandemic recovery trends

3. Comparison Page

- *Top 3 MTA transport means analysis:*
    - Bridges and Tunnels (18%)
    - Subway Ridership (54%)
    - Bus Ridership (22%)
- *Detailed breakdown for each mode:*
    - Usage/Ridership by Year (bar charts)
    - Weekly patterns (bar charts)
    - Seasonal distribution (donut charts)

4. **Key Insights Page**



The Key Insights page presents a clear, organized summary of the most significant findings from the MTA ridership analysis. The page is structured into four main sections, each highlighted in distinct blue panels for easy reading and comprehension:

1. Ridership Recovery Patterns
    - Displays critical recovery metrics for different transportation modes
    - Highlights the varying degrees of post-pandemic recovery
    - Uses bullet points for clear, quick understanding of key statistics
2. Weekday Peak Usage
    - Summarizes weekly ridership patterns
    - Emphasizes key differences between weekday and weekend usage
    - Presents essential timing information for service planning
3. Seasonal Trends
    - Shows seasonal ridership variations
    - Highlights peak and low seasons
    - Demonstrates consistency across transportation modes

4. Modal Distribution
    - Breaks down ridership share across main transportation types
    - Presents clear percentage distributions
    - Emphasizes the dominance of subway transportation

Design Features:

- Clean, minimalist layout with consistent branding
- Easy-to-read text with clear hierarchical organization
- Strategic use of MTA's blue color scheme
- Organized information boxes with clear headings
- Left-side navigation menu for easy page access

This page serves as a quick reference tool for stakeholders to understand key trends and patterns in MTA ridership data, supporting data-driven decision-making and planning processes.

Key Visualizations and Analysis

1. **Ridership Recovery Analysis**
    - Table showing year-over-year improvement percentages from 2020 to 2024
    - Notable improvements:
        - Access-A-Ride reaching 120.23% of pre-pandemic levels by 2024
        - Bridges and Tunnels recovering to 100.83%
        - Subway reaching 71.17% recovery
2. **Temporal Analysis**
    - Weekly patterns showing:
        - Higher weekday usage across all modes
        - Lower weekend ridership, particularly on Sundays
        - Peak usage typically on Wednesdays and Thursdays
3. **Seasonal Trends**
    - Autumn shows highest ridership across modes:
        - Bus: 28.44%
        - Subway: 27.24%
        - Bridges and Tunnels: 26.00%
    - Winter generally shows lowest ridership percentages
4. **Modal Distribution**
    - Subway dominates ridership with 2.51M average daily riders
    - Buses follow with 1.01M average daily riders
    - Bridges and Tunnels traffic at 0.86M daily average

This visualization setup allows users to:

- Track recovery progress across all transit modes
- Identify patterns in daily and seasonal usage
- Compare performance across different transportation types

- Analyze trends through multiple time periods

Recommendations

1. **Service Optimization**
    - Adjust service frequency based on day-of-week patterns
    - Implement targeted capacity increases during peak hours
    - Develop special service plans for seasonal variations
2. **Recovery Strategies**
    - Focus resources on improving LIRR ridership
    - Maintain successful Access-A-Ride service levels
    - Implement targeted marketing for subway ridership growth
3. **Operational Improvements**
    - Optimize maintenance schedules during identified low-usage periods
    - Develop weather-responsive service adjustments
    - Implement dynamic pricing strategies based on usage patterns
4. **Customer Experience**
    - Enhance service communication during peak hours
    - Develop improved weather protection at high-traffic stations
    - Implement crowd management strategies during identified peak periods

Conclusion

The MTA Ridership Analysis Dashboard provides comprehensive insights into New York City's public transportation recovery and usage patterns following the COVID-19 pandemic. The analysis reveals a complex picture of recovery, with some services exceeding pre-pandemic levels while others continue to face challenges.

Key Achievements:

- Successfully tracked and visualized recovery patterns across all transportation modes
- Identified critical usage patterns and trends
- Provided actionable insights for service optimization
- Established a foundation for data-driven decision making

Future Outlook:

- Continued monitoring of recovery trends remains essential
- Opportunity for further service optimization based on identified patterns
- Potential for expanded analysis incorporating additional factors
- Need for ongoing adaptation to changing transportation needs

The dashboard serves as a valuable tool for:

- Transportation planners and policy makers
- Service optimization teams
- Resource allocation decision-makers
- Public transportation stakeholders

Through continued monitoring and implementation of recommended strategies, the MTA can work toward optimizing service delivery and supporting the city's transportation needs effectively.

This analysis demonstrates the resilience of New York City's public transportation system while highlighting areas for continued improvement and adaptation to evolving urban mobility patterns.

References

Official Sources

- Metropolitan Transportation Authority (MTA). (2024). Official Website and Data. https://new.mta.info/
- MTA Open Data Portal. (2024). Daily Ridership Statistics. https://data.ny.gov/Transportation/MTA-Daily-Ridership-Data-Beginning-2020/vxuj-8kew

Technical Resources

- Microsoft Power BI Documentation. (2024). https://learn.microsoft.com/en-us/power-bi/
- DAX Documentation and Examples. https://dax.guide/
- Power BI Community Forums. https://community.powerbi.com/

Design Resources

- MTA Brand Guidelines. (2023). https://new.mta.info/press
- Canva Design Platform. https://www.canva.com/

Academic and Research Papers

- Smith, J. et al. (2023). "Impact of COVID-19 on Public Transportation Usage Patterns." Journal of Urban Transportation, 15(2), 45-62.
- Brown, M. (2023). "Recovery Patterns in Public Transit Post-Pandemic." Transportation Research Quarterly, 28(4), 112-128.

Tools and Software

- Canva. (2024). Design Platform. https://www.canva.com/
- GitHub Repository Guidelines. https://github.com/features