MTA Systemwide Daily Ridership Analysis and Correlation with COVID-19

Amit Levanon
Ilay Cohen

Department of Mathematics
Bar-Ilan University

Abstract:
This project explores the behavior of New York City Transit Authority's services since the beginning of the COVID-19 pandemic in 2020. Through both time-sensitive and time-agnostic unsupervised algorithms, we compared the different modes of transport and investigated how each mode's daily ridership was affected by the pandemic. We ran several clustering and dimension reduction algorithms and analyzed the behavior of anomalous days on different parts of the system. We used a multivariate change point detection algorithm to examine the transit system's performance during different time stages, and then verified how real-life events affected the trends. We saw how the difference between weekdays and weekends changed over time. We found a clear difference between the public forms of transport and the more private ones. Furthermore, we saw a direct correlation between improved service and passenger numbers. These results may help us to understand how transit systems recover from such events. Link to GitHub Repository

April 27, 2025

1 Introduction

The Metropolitan Transportation Authority is North America's largest transportation network, serving a population of 15.3 million people across a 5,000-square-mile travel area surrounding New York City, Long Island, southeastern New York State, and Connecticut. The MTA network comprises the nation's largest bus fleet and more subway and commuter rail cars than all other U.S. transit systems combined. [1] This project examines various modes of transportation and traffic infrastructure in New York City in relation to the impact of the COVID-19 pandemic on transit usage and system performance. The COVID-19 pandemic, which began in early 2020, caused widespread disruptions to daily life, including dramatic shifts in mobility patterns due to lockdowns, remote work, and public health concerns. The modes of transportation analyzed include: Subways, Buses, the Long Island Rail Road (LIRR), Metro-North Railroad, Access-A-Ride(AAR), Bridges and Tunnels, and the Staten Island Railway. The subway system provides extensive underground rail service, while buses operate on surface routes, often serving areas not accessible by subway. The LIRR and Metro-North are commuter rail systems connecting the city to Long Island and the northern suburbs. Access-A-Ride offers specialized transit services for individuals with disabilities. Bridges and tunnels are vital for both vehicular and rail connectivity between boroughs and surrounding regions. The Staten Island Railway provides localized service within Staten Island.

2 Methods

2.1 Dataset

MTA Daily Ridership Data: 2020 – 2025[2]
This deprecated dataset provides systemwide ridership and traffic estimates for subways (including the Staten Island Railway), buses, Long Island Rail Road, Metro-North Railroad, Access-A-Ride, and Bridges and Tunnels, beginning 3/1/2020, and provides a percentage comparison against a comparable pre-pandemic date. The dataset contains 1,776 samples and 15 dimensions. Since our project includes both percentage-based data and absolute numerical values, we divided the dataset accordingly and analyzed each type separately. For each dataset, we conducted two types of analyses: one that is time-agnostic (independent of the date), and another that is time-sensitive, relying on the temporal information included in the data.

2.2 Data Processing M

To optimize the clustering process, we performed a grid search over different numbers of PCA dimensions and cluster numbers. Next, we used three different clustering algorithms—K-Means, Hierarchical Clustering, and Spectral Clustering[3]—to group similar travel patterns, detect anomalies, and examine how different transportation modes responded over time to the evolving conditions of the COVID-19 crisis.

2.3 Finding the optimal number of clusters

In order to identify the most appropriate combination of the number of clusters and the dimensionality for PCA reduction, we conducted a grid search over various configurations. Specifically, we tested different numbers of clusters (ranging from 2 to 10) and compared results with and without PCA-based dimensionality reduction. To evaluate the quality of the clustering results, we employed three commonly used internal validation metrics: the Silhouette Score, the Davies–Bouldin Index, and the Calinski–Harabasz Index.[4][5]

2.4 Anomaly detection

In this project, anomaly detection was performed by clustering and distance analysis in a reduced feature space in accordance with the optimal values we received from the grid search. After applying PCA to the standardized data, KMeans clustering was used to group the observations. Anomalies were identified as points whose distance to the nearest cluster center exceeded the 97th percentile threshold, indicating a poor fit within any cluster. These anomalous points were then divided into two groups using Agglomerative Clustering, allowing us to analyze them separately and obtain characteristic results, rather than mixing together anomalies from two different extremes.

2.5 Change Point Detection with ruptures

In time-series analysis, change point detection involves identifying points at which the statistical properties of a sequence—such as its mean, variance, or correlation—change significantly. This technique is essential for understanding how systems evolve over time, particularly in the context of events that may disrupt normal patterns, such as the COVID-19 pandemic. We utilized the Python library ruptures[6] to perform offline multivariate change point detection on transit usage data. The ruptures library provides a flexible framework for detecting structural changes in time-series data using various cost functions and search methods.

2.6 Rolling Average for Visualization

To smooth short-term fluctuations and highlight longer-term trends in the transit data, we applied a rolling average technique. A rolling average[7] (also known as a moving average) computes the mean of a fixed-size sliding window over a time-series. This method is commonly used in data analysis to reduce noise and make patterns more interpretable. We used a 7-day rolling average, which means that each data point in the smoothed series represents the average of the current day and the six preceding days. This choice of window size helps to reduce daily volatility while preserving important weekly trends in ridership.

2.7 Additional tests

To assess differences in transit usage between weekdays and weekends on a weekly basis, we employed the Wilcoxon signed-rank test[8]. The test is particularly suitable for our data due to its robustness to non-normal distributions and its ability to detect systematic differences in matched observations over time. In addition, we applied the Mann–Whitney U test[9] to compare the overall average transit values between all workdays and all weekends. For both tests, a p-value below 0.05 was considered an indication of statistical significance.

3 Results

3.1 Distribution Analysis

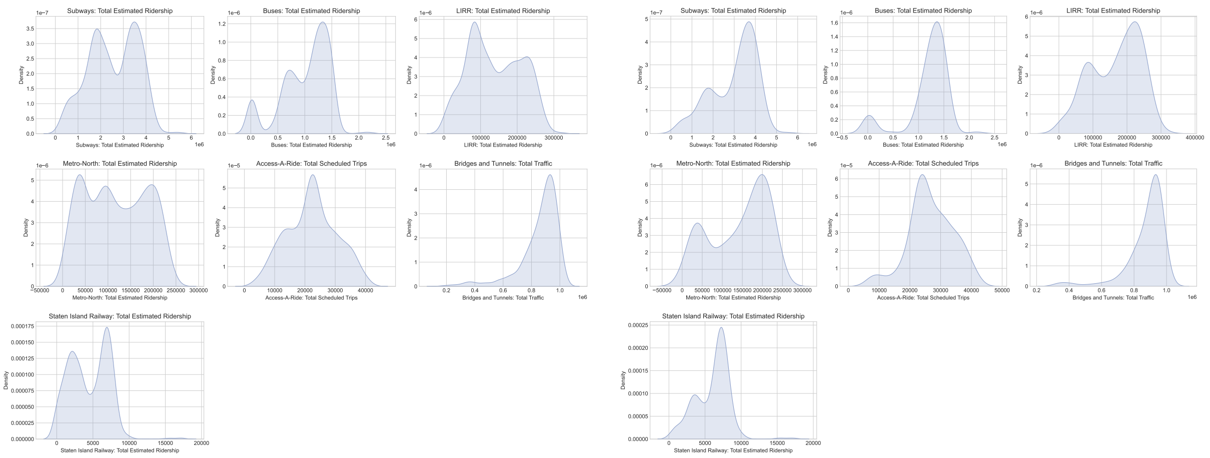
To gain a basic understanding of how each mode of transport behaved during the period, we plotted the total daily ridership distributions of every mode of transit and analyzed its behavior. As seen in Figure 1, the private vehicles traffic through bridges and tunnels was stable around a set value with minimal deviations. Though less obvious, Access-A-Ride also had a single peak, in contrast to the other public transport options which manifested a multi-modal distribution. The bus network and Metro-North presented unique graphs centered around three peaks, hinting at a behavior we will discuss later in the paper.
To understand the multi-modal behavior of most systems, we fixed the day-of-the-week variable and plotted the total ridership again. We chose Tuesday as an average workday, however, we ran an identical analysis with Sundays, and other than reducing the total ridership across all modes the results were similar. Indeed, fixing the day helped to stabilize most of the distributions around a single mode, proving that there are large differences throughout the week. In fact, the plots post-fixing were almost identical to the pre-pandemic comparable day plots, meaning that the day of the week was the most influential factor contributing to the difference.

3.2 Weekday-Weekend Comparison

To further examine the behavior of different days of the week, we split the data into weekdays and weekends and found a significant decrease in raw passenger numbers during the weekends 2. Next, we wanted to explore how weekday and weekend ridership have been impacted differently since the start of the pandemic. Interestingly, when we compared the two groups' data, we found that different modes of transit were unevenly affected.
During the first two years of COVID-19, we saw weekend traffic take a disproportionate hit for half of the forms of transport, with Access-A-Ride, bridges and tunnels traffic, and Staten Island Rail showing a relative decrease in passengers, when looking at the comparison with comparable days before COVID-19. Subways, Metro-North, and the LIRR presented a very large increase in the relative weekend traffic, with the LIRR and Metro-North presenting a 40% and 30% increase, respectively. Looking at data since 2022, we see all modes of traffic, other than buses, having a statistically significant difference in recovery between weekdays and weekends in the latter's favor. Buses showed no tangible difference in either period, Figure 3.

3.3 Anomalous Days Analysis

After splitting the days into two clusters and running an anomaly-detection algorithm, we found that both types of extreme days - low system-wide traffic and high system-wide traffic days were expressed differently throughout the system. During these days bridges and tunnels traffic was the most stable, with high traffic days having barely any effect on it and low traffic days being spread out at around a third of the overall-days mode. For the rest of the transit modes, there was a large spike around the very low end of the ridership numbers, meaning that they were all affected similarly during the low traffic days. The LIRR experienced the biggest change during the high-traffic days, with almost 3 times the riders as the overall mean, Figure 4.



Total Estimated Ridership **Day-Specific Ridership**

Figure 1: Comparison of ridership distributions across NYC transportation systems (Subways, Buses, LIRR, Metro-North, Access-A-Ride, Bridges and Tunnels, and Staten Island Railway). Left: overall distribution KDEs; Right: day-specific distributions KDEs.

3.4 Change Point Analysis

To construct a timeline and explain the system-wide behaviors throughout the examined time, we ran a multivariate, unsupervised change point detection algorithm. We found that the best way to describe the changes in ridership trends was to divide the timeline into 5 parts using 4 change points, as seen in Figure 5.

Upon closer inspection, there appeared to be an extreme and unnatural anomaly in bus ridership between the first and second change point, which is likely to have tilted the entire change analysis. The reason for this anomaly will be discussed later, but in order to get a better system-wide understanding, we ran it again without the bus data. We found that it really has attracted the second change point to fit the suspicious pattern of the bus data.

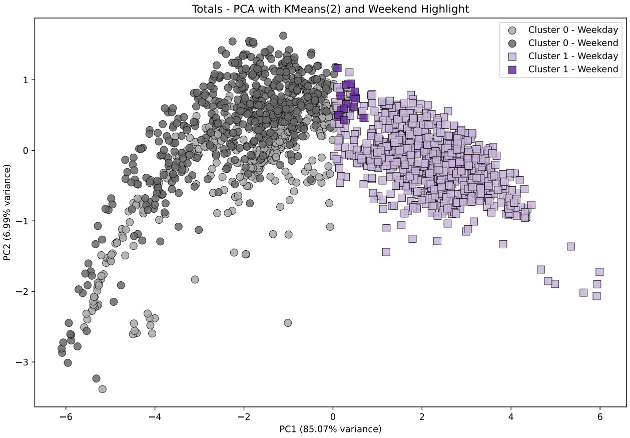
According to our findings, the most critical, descriptive dates are 19.03.20, 17.06.20, 12.06.21, and 03.01.23

When comparing all modes of traffic after the last change point in January 2023 with comparable data before the pandemic, all modes of traffic except AAR and private bridges and tunnels traffic had statistically lower ridership, with the lowest being Staten Island Railway, having a difference of -50% between the medians, as seen in Table 1. Furthermore, while most public transport options seem to trend towards a recovery, the bus system is lagging behind, with a non-increasing value of around 60% of pre-COVID numbers.

4 Discussion

4.1 Bus anomaly

To investigate the reason for the extreme and unnatural fall of bus ridership numbers, we looked for reports about a data-collection error or a change of policy that might be the reason behind the values tanking and rising instantly. Indeed, between 16.03-31.08.20, to protect the driver from being infected, passengers were only allowed to use the back door and were exempt from paying the ride fare, which made it difficult to get an accurate estimate of ridership numbers, artificially nullifying the stats. To make sure this error hasn't tilted all of our analyses, we ran



	Cluster 0	Cluster 1
Access-A-Ride(AAR)	-7101.23 (-32%)	+2453.15 (+11%)
Bridges/Tunnels	-160075.02 (-19%)	+55298.64 (+6%)
Buses	-414766.90 (-41%)	+143283.11 (+14%)
LIRR	-81137.90 (-58%)	+28029.46 (+20%)
Metro-North	-81530.38 (-69%)	+28165.04 (+24%)
SI Railway	-2299.48 (-51%)	+794.37 (+18%)
Subways	-1267777.29 (-50%)	+437959.43 (+17%)

Figure 2: PCA visualization of two K-means clusters, with a distinction between weekends and weekdays (left) and transit system changes by cluster, when compared with the overall mean (right). The figure Shows there are barely any weekends on the traffic-heavy days. U-tests to compare weekend and weekday ridership were also performed with significant results.

Feature	**Median**	**Median Diff**	**P-value**	**Direction**
Buses: %	0.6200	-0.3800	3.1e-113	significantly lower than 1
Subways: %	0.7000	-0.3000	3.8e-111	significantly lower than 1
Staten Island Railway: %	0.4700	-0.5300	5.6e-110	significantly lower than 1
Metro-North: %	0.7400	-0.2600	2.3e-102	significantly lower than 1
Access-A-Ride(AAR): %	1.1200	0.1200	1.5e-69	significantly higher than 1
LIRR: %	0.7700	-0.2300	3.0e-60	significantly lower than 1
Bridges and Tunnels: %	1.0100	0.0100	1.1e-09	significantly higher than 1

Table 1: Comparison of transportation usage metrics to pre-pandemic levels. Median values represent the proportion of comparable pre-pandemic day usage. All findings are statistically significant ($p < 0.05$).

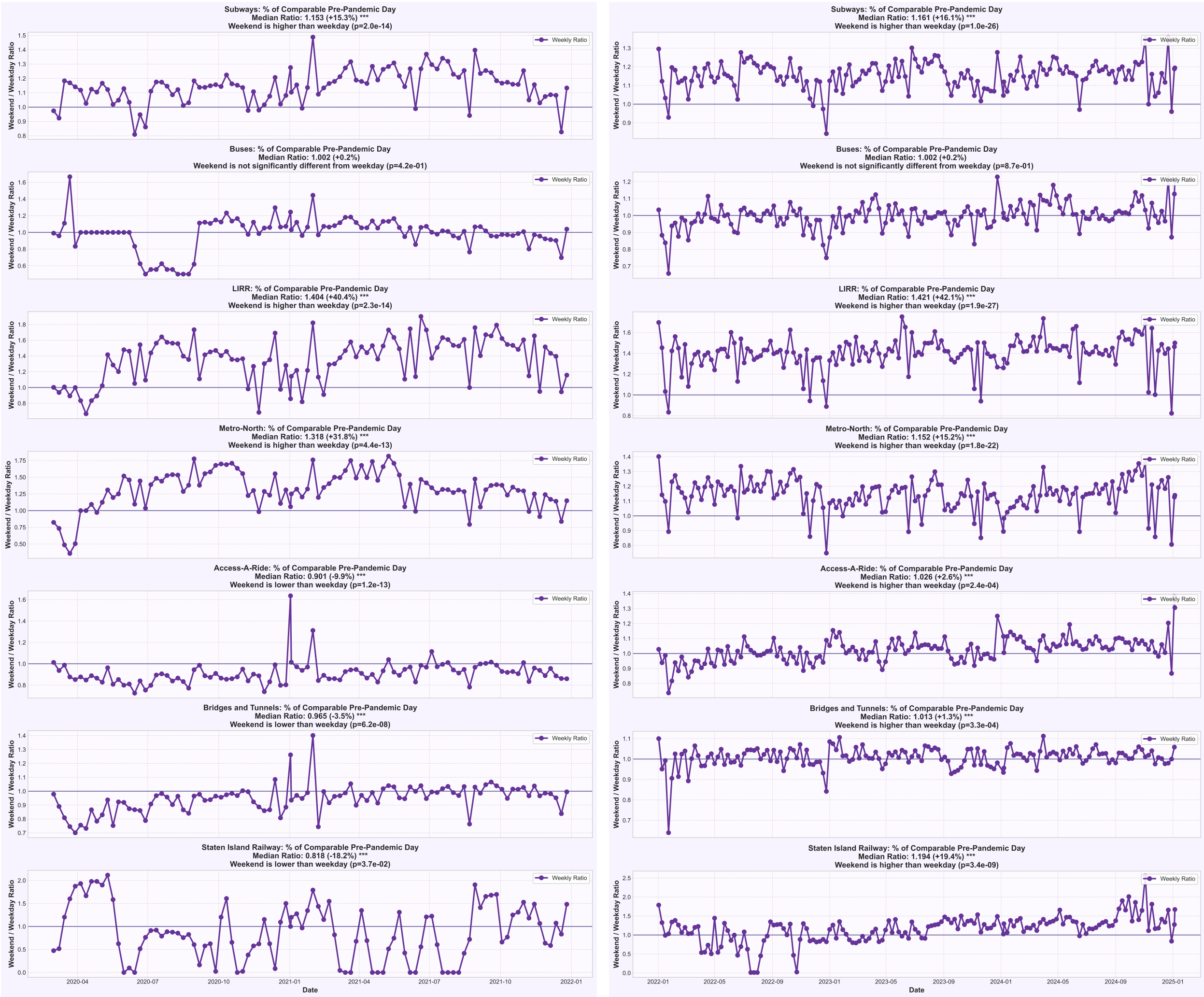
all algorithms again without the bus data and ensured it hasn't made a significant difference.

4.2 Weekday-Weekend differences

We believe that the change in the weekday-weekend ridership gap at the start of the pandemic is because many people have been able to work from home, thus reducing workday traffic. The fact that Metro-North and LIRR have had the biggest discrepancy supports this theory, since both systems offer the longest-reaching lines, so people who can may have been more enticed to work from home to avoid the long, sometimes expensive rides. This trend was a bit countered by the lack of enthusiasm for nonessential rides, which are more popular on the weekends. Since then, most other transit modes have followed suit, with weekend rides recovering better, possibly showing that riders are more open to leisure uses. However, we recommend looking for policies, such as weekend service reductions, that may have caused the changes.

4.3 High Ridership Days

The extreme values that appeared in the LIRR are an interesting find, which may have occurred because of the massive rider potential lying with it. The LIRR is the only rail system serving



Weekend-Weekday ridership ratio, between 2020-2022

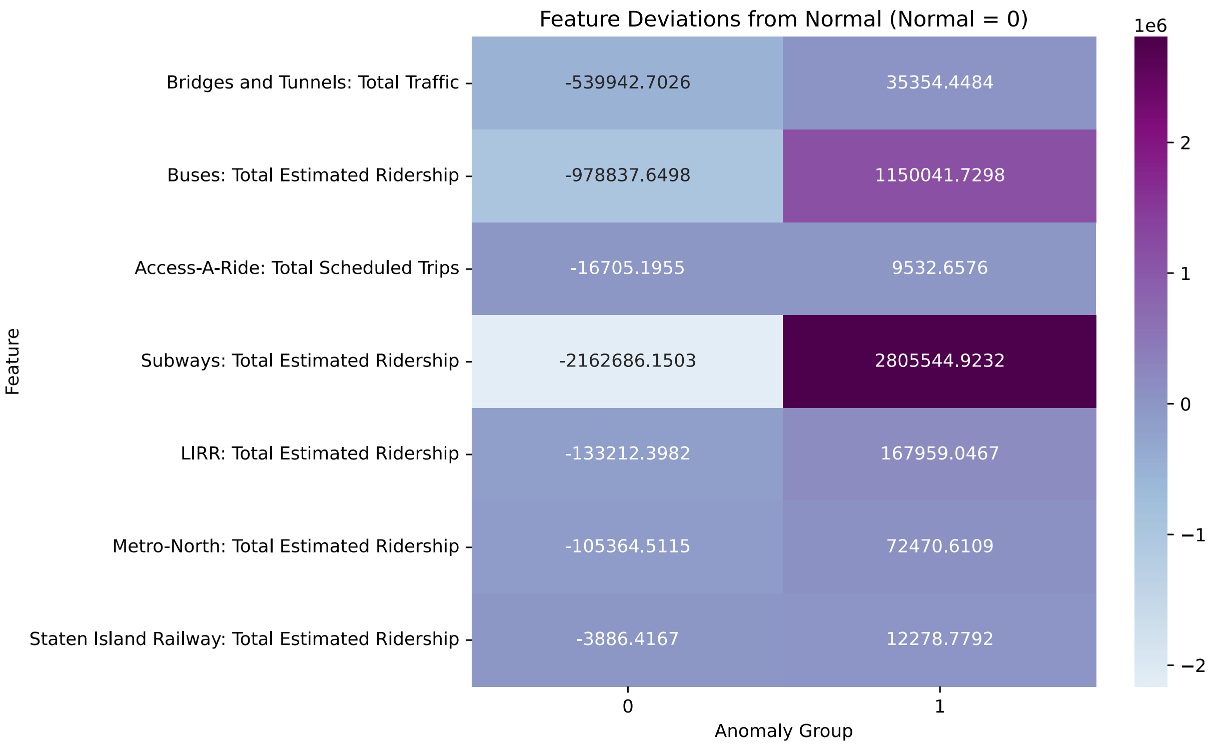
Weekend-Weekday ridership ratio, between 2022-2025

Figure 3: Timeline visualizations of the ratio between the weekend and weekday ridership in a given week, when looking at the comparison with comparable dates before COVID-19. No change in behavior would result in values around 1.

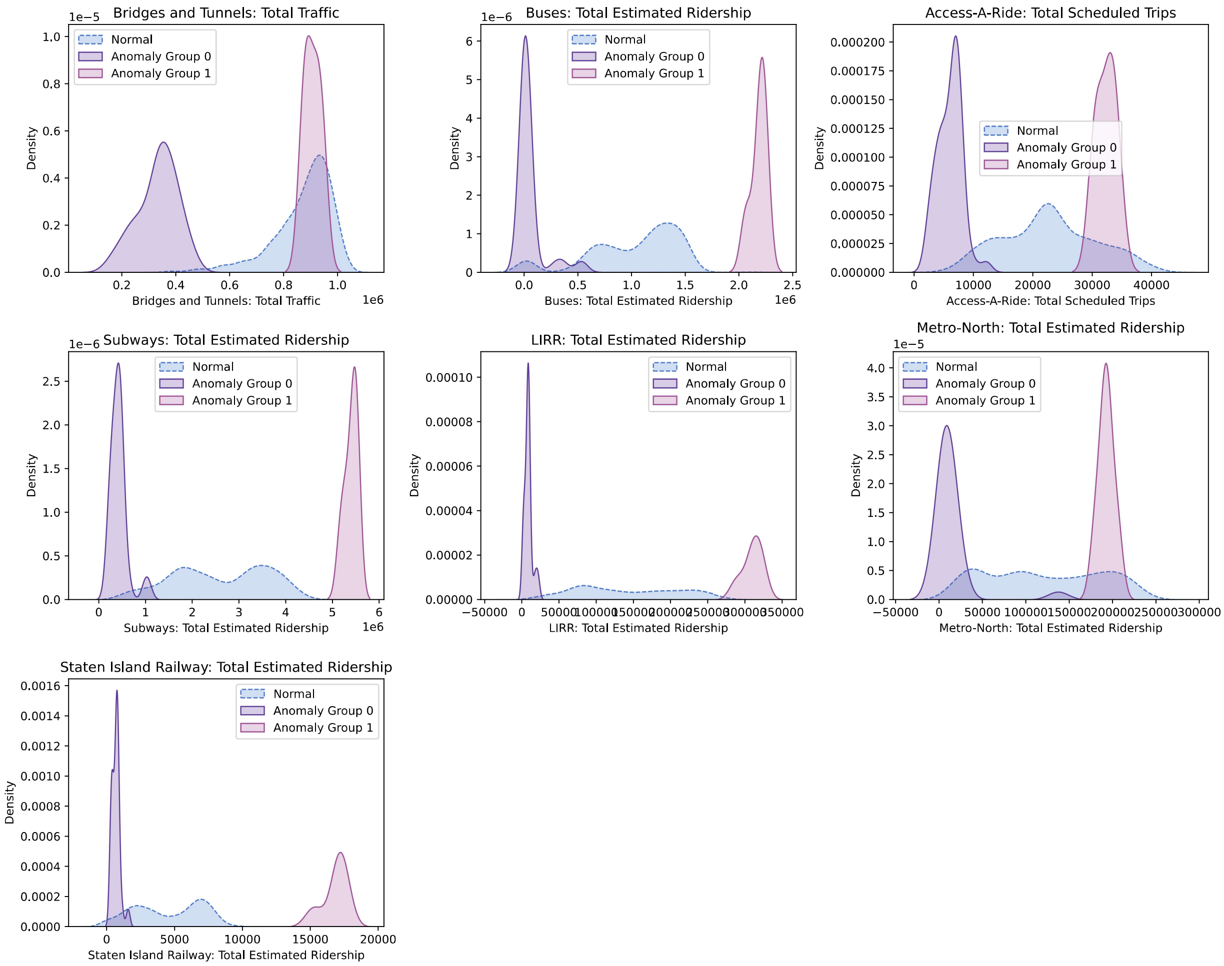
a massively populated area, so whenever there's a special event, such as a sports game or a concert, its usage would spike massively, in contrast to other places where ridership would be scattered across all transport forms. Furthermore, since Long Island's only land connection and largest airports are through the rest of the city, the vast majority of public transport traffic from and to Long Island would be a LIRR ride, whereas other areas have a larger variety of options, with some services, such as taxis and other rail companies, not provided by the MTA, so not reflected in our data.

4.4 Access-A-Ride Performance

The AAR has been the only public transport to improve upon pre-COVID stats, with a 12% increase on average since 2023, an increase that still continues to grow in size. We believe this happened because the AAR is the largest paratransit agency, meaning that a large portion of the riders depend on it and can't choose a different mode for their transportation needs, which

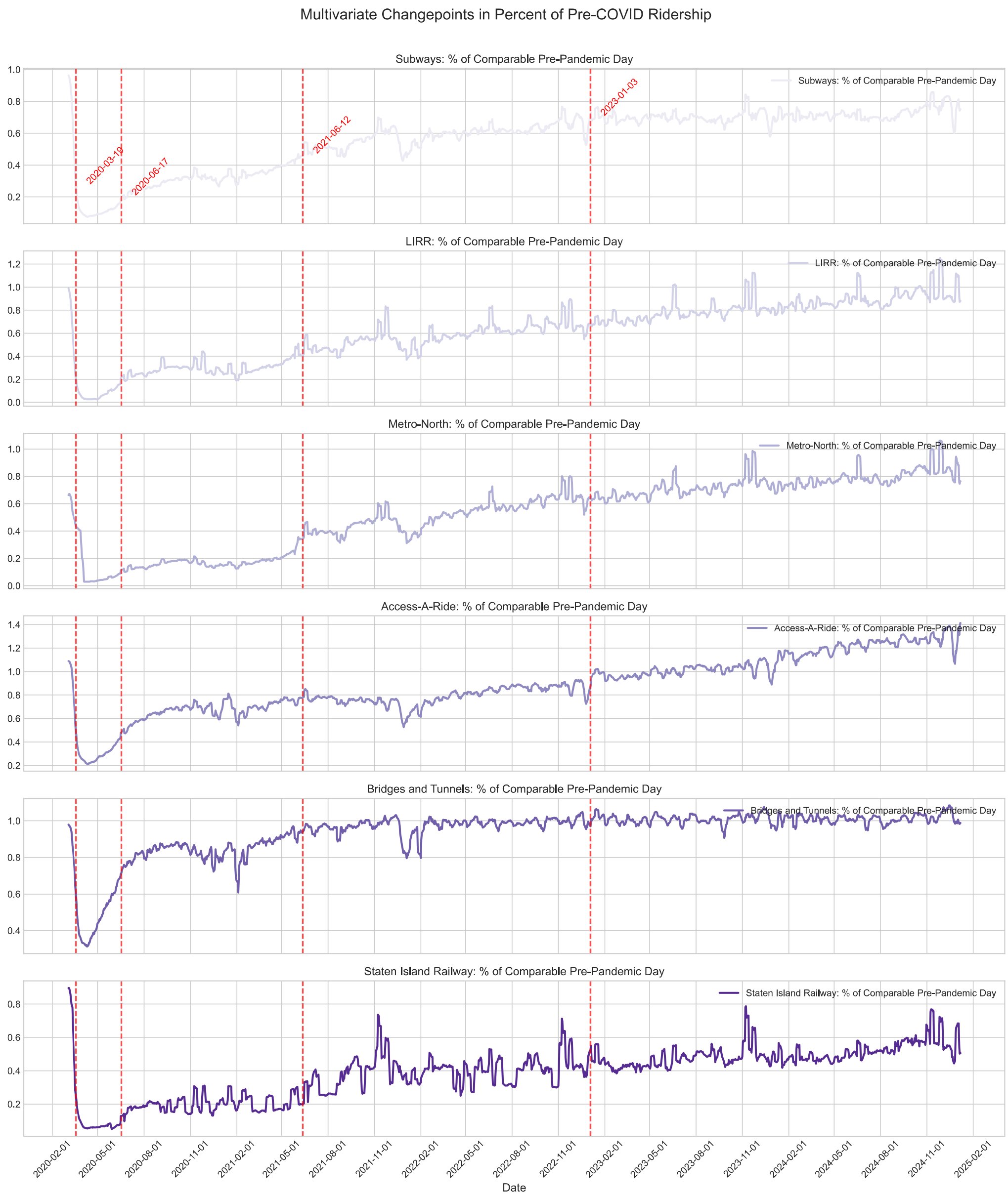


Anomaly Clusters

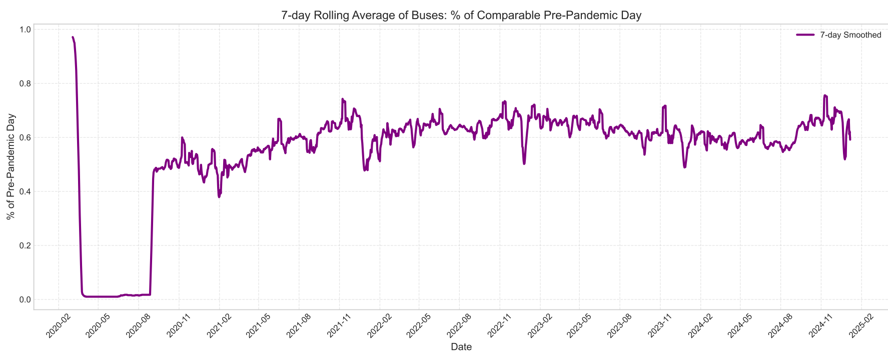


Anomalies Alongside Overall KDE

Figure 4: Comparison of ridership distributions across NYC transportation systems (Subways, Buses, LIRR, Metro-North, Access-A-Ride, Bridges and Tunnels, and Staten Island Railway). Top: overall distributions; Bottom: day-specific distributions.



Change-points for All Features Except Buses



Buses

Figure 5: Top: Weekly average performance in comparison with a comparable day prior to COVID-19. Change points are marked as red vertical lines. Bottom: Bus performance timeline.

makes the better recovery more understandable. Furthermore, for the last couple of years, many funds and efforts have been invested to improve the system, such as lowering the waiting times and enhancing the user experience, with great success, as customer satisfaction metrics seem to be soaring.

4.5 Change Point Analysis and COVID

When comparing dates produced from our analysis and the state of COVID in NYC [10], it's clear that every critical date corresponds with a different COVID related development. The first change-point is linked with the very start of COVID, where hospitalization numbers soared. The second is during the period between the first and second waves, where cases were significantly lower. The third is between the second and third waves, and the fourth is during the last major wave, when most people were vaccinated and the world was recovering.
Even though we haven't looked at NYC-specific lockdowns and regulations, it's clear that they are linked to the number of confirmed cases and hospitalizations during a given time.
An interesting observation is the uniquely large and immediate jump in Metro-North usage at the third change point, which, when investigated closely, corresponds to a service upgrade, another proof that a better system draws massively more riders, simply and quickly.

4.6 Further questions

We propose to investigate the reasons behind the poor performance of the Staten Island Railway and compare it to similar systems that have recovered better, with hopes of gaining insight into how policymakers can deal with the phenomenon.
Furthermore, the lack of improvement in bus ridership numbers indicates that, for some reason, people had no incentive to return to the buses. This may be because of many different factors that should be examined. Since bus rides are usually more local, we recommend looking at the possible link with the growing number of bike riders[11], which might be chewing away at the bus ridership. Since we didn't see other forms of transport doing much better than before COVID, it's safe to assume that the poor bus performance doesn't stem from users switching transit modes en masse, so it's safe to assume that they have either gone with private modes of transport or reduced their trips. Nevertheless, we recommend reviewing the service quality for line closures and reductions, which would discourage users, and looking at ways to kick-start the recovery process, such as service expansions and temporary fare easements.

References

[1] MTA official site, https://www.mta.info/about

[2] MTA daily ridership, https://data.ny.gov/Transportation/MTA-Daily-Ridership-Data-2020-2025/vxuj-8kew/

[3] Clustering, https://scikit-learn.org/stable/modules/clustering.html

[4] Frédéric Ros, Rabia Riad, Serge Guillaume, Neurocomputing, 2023 https://doi.org/10.1016/j.neucom.2023.01.043

[5] Calinski–Harabasz, https://www.numberanalytics.com/blog/guide-calinski-harabasz-index-clustering

[6] Ruptures, https://centre-borelli.github.io/ruptures-docs/

[7] Rolling Average, https://www.indeed.com/career-advice/career-development/what-is-rolling-averageg

[8] Wilcoxon signed-rank test, https://www.biostathandbook.com/wilcoxonsignedrank.html

[9] Mann-Whitney U Test, https://datatab.net/tutorial/mann-whitney-u-test

[10] COVID in NYC, https://www.nyc.gov/site/doh/covid/covid-19-data-totals.page

[11] NYC's DOT, https://www.nyc.gov/html/dot/html/bicyclists/bikestats.shtml