COVID-19 and NYC Public Transit Ridership

Drew Brock

December 2022

1 Introduction

The COVID-19 pandemic has improved over the past months, but it is still not over. Because of this it is important to use the data from the past two and a half years to better understand the virus and also to prepare for pandemics in the future. During the pandemic many people's lives were dramatically changed, both socially and professionally. For instance, many people who lived outside the city would use public transportation to commute into the city for work. However, as people started working from home at record rates, many people stopped riding public transportation. In urban areas, public transportation was identified as a major contributor to the spread of COVID-19, and efforts to maintain its safety for those who depend on it have been hampered by various challenges and difficulties. When looking at these factors we can hypothesize that ridership for public transportation likely decreased, but we wanted to take a deeper look into just how much the pandemic and ridership were correlated and if there were any specific components that contributed. We decided to look into the relationship between COVID-19 cases, hospitalizations, and deaths in the New York City metro area and Metro Transit Authority (MTA) ridership.

2 Methods

The primary desired output of the COVID-19 data and the MTA ridership data is one correlation matrix. Before this could happen however, the data needed to be cleaned. The ridership data was pulled from the publicly available and frequently updated NY Gov website in a csv file. The COVID-19 data was pulled from the NYC Department of Health and Mental Hygiene GitHub repository which is updated daily. The chosen data to analyze was the total of subway and bus ridership of the five boroughs correlated to COVID-19 cases, hospitalizations, and deaths over time. Since the data is a time series, it was re-sampled over the daily numbers from 03/01/2020 - 11/20/22. After this point all that needs to be done is to put the data together and correlate it.

3 Results

The results of this are quite interesting. The matrix produced shows a couple notable things.

	Total	*Hospitalized*	*Deaths*	*Cases*
Total	1.000000	−0.303980	−0.414366	0.026873
Hospitalized	−0.303980	1.000000	0.771436	0.544455
Deaths	−0.414366	0.771436	1.000000	0.153147
Cases	0.026873	0.544455	0.153147	1.000000

Figure 1: Correlation matrix of total subway and bus ridership, hospitalizations due to COVID-19, deaths due to COVID-19, and cases of COVID-19.

First, there is a slightly notable negative correlation between hospitalizations and ridership as well as deaths and ridership. It can be interpreted from this information that as more people rode transit over the course of the pandemic, it was more likely that less people were being hospitalized or dying due to the pandemic. If people feel more confident about the state of things, the more likely they are to ride transit.

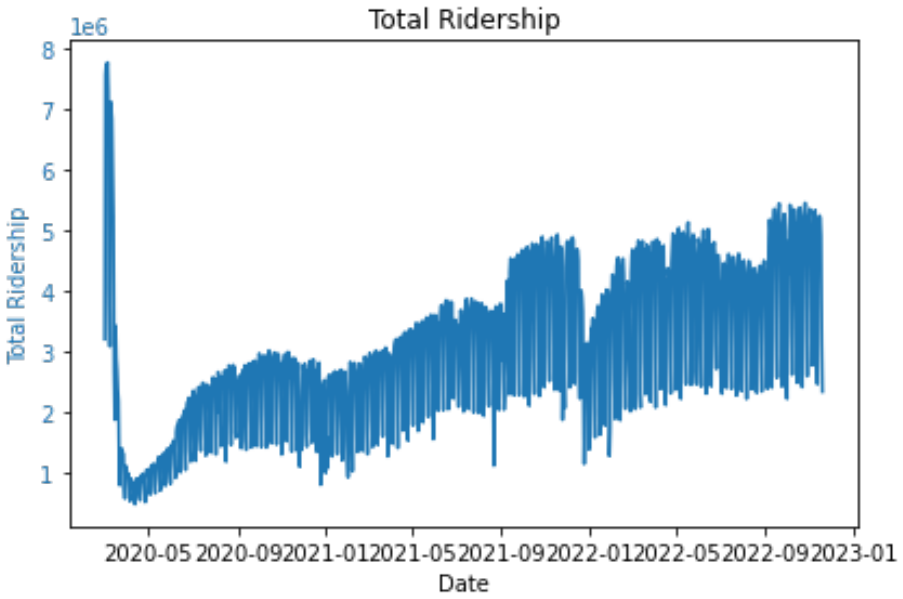

Figure 2: Total Ridership on subways and buses from 03/01/2020 - 11/20/2022.

The next correlation that should seem quite obvious is that cases of COVID-19 are positively correlated with hospitalizations, but weakly correlated with

total deaths. This makes sense as more often than not people recover from the illness and the worst case scenario is most likely being hospitalized. However, a sobering truth is that hospitalizations are strongly correlated with deaths. Finally, one can see there is a noticeable dip at the beginning of January 2022.

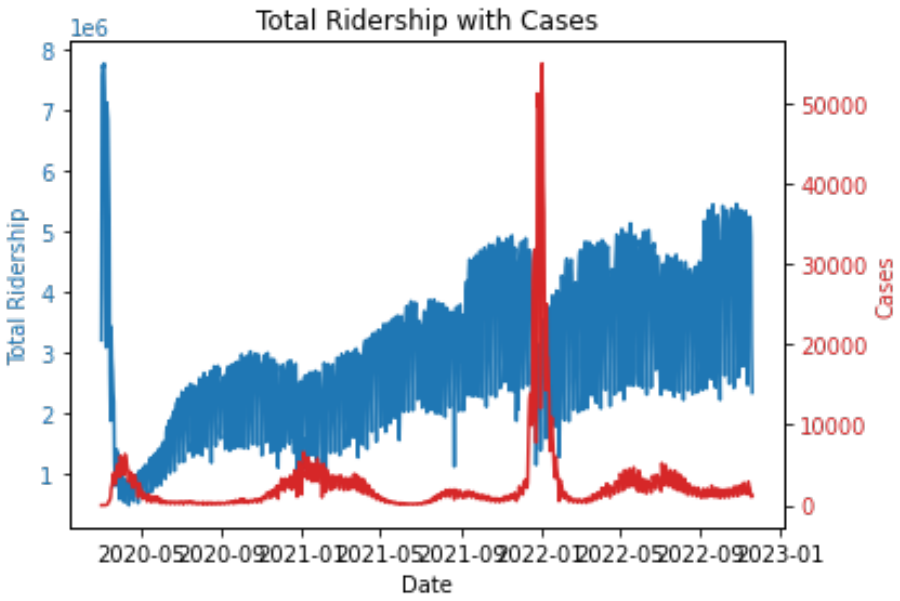

Figure 3: Total ridership on subways and buses overlaid with cases of COVID-19 from 03/01/2020 - 11/20/2022.

Quite clearly this is the impact of the omicron variant. The steep drop of cases did not lead to a staggering rise in hospitalizations or deaths. Impressively, ridership went back to pre-variant within less than 5 months. Overall the correlation between COVID-19 and public transit is more so related to hospitalizations and deaths rather than just number of cases.

4 Conclusion

Our analysis shows that the number of hospitalizations and deaths due to COVID-19 is strongly correlated, with more hospitalizations leading to more deaths. There is also a weaker but still significant correlation between the number of cases and hospitalizations. When it comes to the relationship between ridership on the New York City subway and bus systems and the number of COVID-19 cases, deaths, and hospitalizations, there is a negative correlation. This suggests that as deaths and hospitalizations decreases, ridership increases. The actual factors behind this are not fully known. Public transit can generally

be considered inelastic as it is such a core part of moving people to places they want and need to be, and the pandemic destroyed that consideration. Following the trends of this analysis, ridership will most likely continue to grow as the pandemic lessens its grip on humanity.

5 Links to Data Sets

`https://data.ny.gov/Transportation/MTA-Daily-Ridership-Data-Beginning-2020/vxuj-8kew`
`https://github.com/nychealth/coronavirus-data`