Understanding the Effects of the Pandemic and Infrastructure Improvement on the Usage of the L-line in Brooklyn NY

Mallory Banks, METIS: Exploratory Data Analysis Project 1

## Introduction: A Study of MTA Usage of the L-Line

As discussed during the selection process, with the pandemic slowing in New York City, we are looking to leave a robust Urban planning guide for the new incoming Mayor of New York for the Borough of Brooklyn. With the recent finishing of the L train renovations and it's positive impact on the transportation trends between the boroughs of Brooklyn and Manhattan we are looking to recommend to the newest Mayor's office places where improving, widening, increasing the efficiency of, or even adding new subway stations may be of great benefit to the people of Brooklyn.

What kind of insight can you provide on our recommendations on strategic investment in transportation? Should we recommend taxpayer dollars for more subway access?

Feel free to let us know if there are further clarifying questions for us and we look forward to the Project proposal in the coming days.

Best.

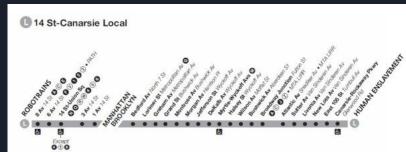
The Mayor's office of:

Bill de Blasio

# Bill de Blasis

### **Problem Statement and Objectives:**

With both the recent completion of construction on the Stations along the L - Line connection Brooklyn to Manhattan as well as the ongoing pandemic, it is imperative that the Mayor's office understand the ROI of the multi-million dollar project as well as right size their recommendation to the incoming mayor's office on continued investment in subway infrastructure in King's County. **The goal of this analysis is to answer:** Is the L-line more in use now than before as a function of construction? Where, if any station, should there be further investment in infrastructure due to high levels of usage?



## Methodology

### Data:

This data set is publically available through the MTA data site for the city of New York. This dataset contains data on turnstiles for all stations of New York including entry and exit data as well as Date, station, and line name. For the purposes of this project, this dataset was subsetted on only the L line for the specified time period of Q2 2019, 2020, and 2021.

### Tools:

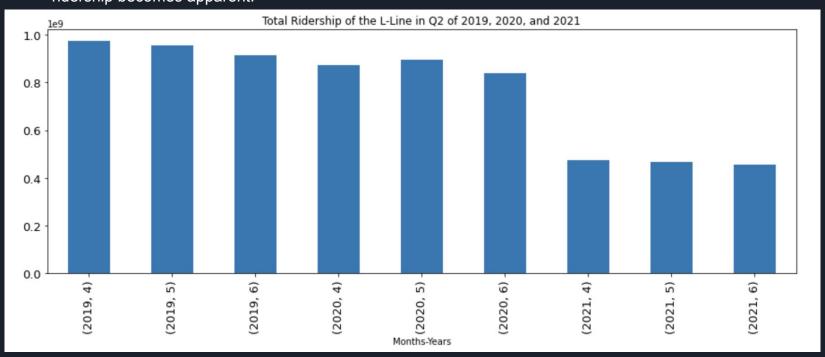
 Tools used included sqlite3, jupyter notebooks running Python and using packages pandas, matplotlib, and seaborn.

### **Metrics:**

 Metrics used were primarily total and average daily entries for both turnstiles and data across station, months, years, and week days

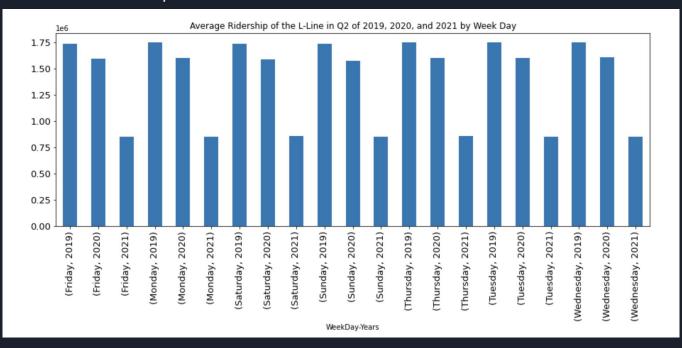
# **Data Suggests that L-Line Usage has Fallen YoY**

When looking at the graph for average daily use of the L -Line Year over Year for Q2, the drop in ridership becomes apparent:

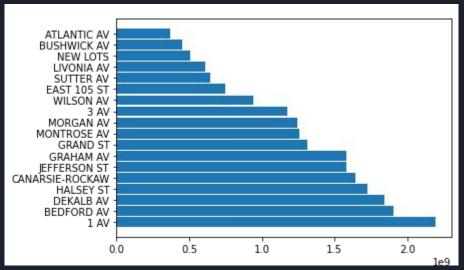


# Average Daily Usage Looks Similar Across Days of the Week

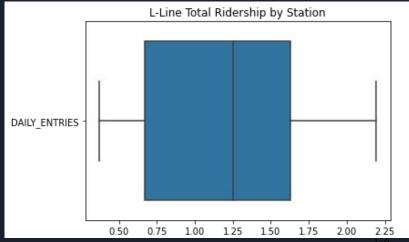
This suggests that the users of the L-line are those who may be dependent on it or are regular users of this train with little spikes seen for the weekend



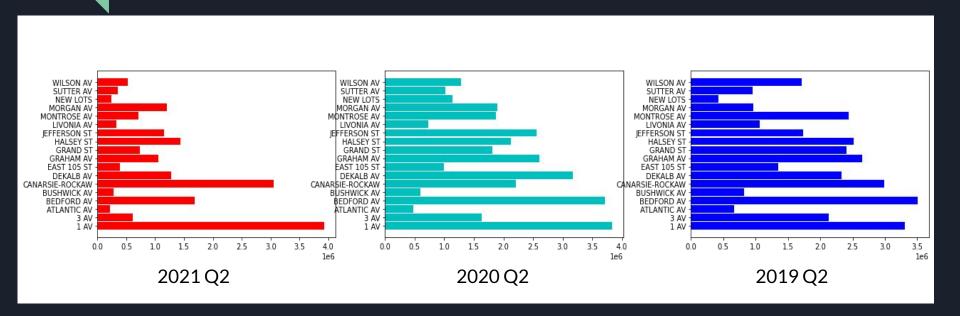
# When analyzing most popular stations, we see clear shifts in patterns YoY



This bar graph represents the most popular L-Line stations in descending order Aggregated across Q2 of 2019,2020, and 2021. This Box plot represents the Distribution of the Data Across Station. Each station above Grand St is in high use. Those Below Sutter is in low popularity.

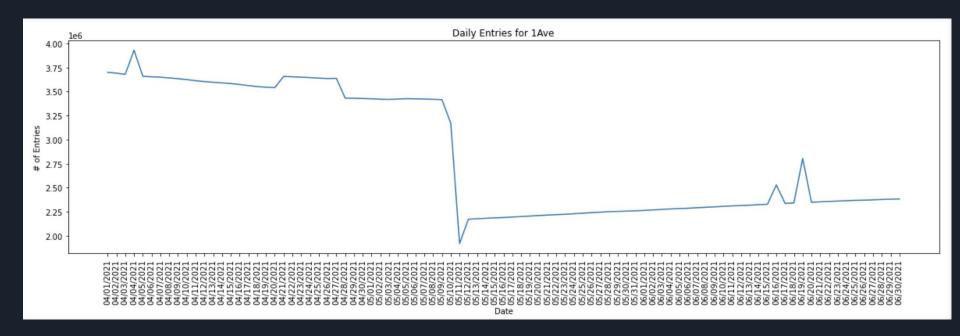


# When analyzing most popular stations, we see clear shifts in patterns YoY Cont.



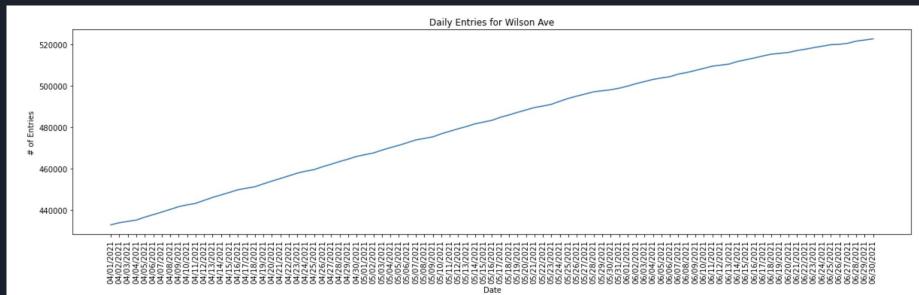
## A Closer Look: 1 Ave

When looking at Daily Entries for the most popular stop on the L-Line, we can see that there was a pop in popularity during the announcement of the construction being complete. This has waned, suggesting the construction had limited influence on the popularity of the line overall in the most populous areas.



## A Closer Look: Wilson Ave

However, the analysis is even more interesting when you look at one of the "lesser" used stations that hovered around the ~25th percentile for Daily entries overall in Q2 of 2021.. Wilson Ave has seen a sky rocket in Daily entries after the conclusion of construction last year and the wider availability of the vaccine beginning in March 2021. This may suggest that construction and vaccination had a smaller impact on "more developed" areas like 1st ave and a larger impact on areas in less gentrified areas (like Wilson ave in Bushwick)



### Conclusions

#### **Conclusions**

Concluding the EDA portion of the Analysis here, the data suggests that the completed construction of the L line following damage from Hurricane Sandy has not increased usage of the Line overall. However, the data also suggests usage may be climbing in "less developed" areas where L train access was sorely needed. Additionally, construction was finished in conjunction with the COVID-19 pandemic which has affected MTA ridership city wide and could be more a function of the pandemic changing the way riders are engaging with public transportation than the L-line itself.

#### Recommendations

- When looking for future investment along this line, the 1 Ave Stop is ripe for continuous improvement as it's ridership has remained steady. A close follow up would be the Canarsie Rockaway stop which has increased usage in the difficult year of 2021. Of special note is the Montrose and Wilson Avenue stops which see large decreases year over year but have seen increases recently

### **Future Work**

- While it may be surprising to see the dip in usage of the L line, we suspect that this may be a phenomenon that is not exclusive to the L line only. It is important to understand the context of the usage of the MTA writ large to understand how precipitous this drop truly is. We recommend a comparison of drop in usage as a function of drop in usage in lines all over the city
- We know that in the age of the pandemic there are other forms of transportation that are on the rise. How can these trends be placed in refeleif when compared to Uber/Lyft usage, Citibike, and Bus ridership?
- If possible, understanding these results in conversation with train times (i.e. when are trains arriving and leaving the station) and how do lengthening times and/or delays affect ridership?