C2SMART

CONNECTED CITIES WITH SMART TRANSPORTATION

A USDOT University Transportation Center

New York University

Rutgers University

University of Washington

University of Texas at El Paso

The City College of New York

UTILIZING SOCIAL MEDIA DATA FOR ESTIMATING TRANSIT PERFORMANCE METRICS IN A PRE- AND POST-COVID-19 WORLD

September 2023

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
4. Utilizing Social Media Data for Estimating Transit Performance Metrics in A Pre- and Post-Covid-19 World		5. Report Date September 2023
		6. Performing Organization Code:
7. Author(s) Camille Kamga, Sandeep Mudigonda, Rodrigue Tchamna, Richard Kish		8. Performing Organization Report No.
9. Performing Organization Name and Address Connected Cities for Smart Mobility towards Accessible and Resilient Transportation Center (C2SMART), 6 Metrotech Center, 4th Floor, NYU Tandon School of Engineering, Brooklyn, NY, 11201, United States		10. Work Unit No.
		11. Contract or Grant No. 69A3551747119
12. Sponsoring Agency Name and Address Office of Research, Development, and Technology Federal Highway Administration 6300 Georgetown Pike McLean, VA 22101-2296		13. Type of Report and Period Final report, 3/1/22-9/30/23
		14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

This study explores the benefits of text and sentiment analysis of Twitter data, as well as the applications of such analysis towards understanding customer sentiment, transit performance, and the relationship between the two. Ultimately the objective of this research is to determine which performance metrics customers most frequently mention and are most passionate about, and where these sentiments are directed when riding with Metropolitan Transportation Authority New York City Transit and how these compare and contrast to the performance indicators reported by the agency. This is carried out by analyzing a year's worth of Twitter data collected from users mentioning their transit experience and recognizing trends regarding sentiment. These trends are then compared with the data reported by MTA in order to find a correlation. The results show that customers mostly tweet negative experiences over positive experiences found with service and are most frequently critical about wait time and delays. These are important factors to consider in a metro as unique as New York's. The data analyzed also gives insight on where the most customer issues occur; most notably the Hudson Line when riding Metro-North Railroad, and the Port Washington Line when riding Long Island Rail Road. This correlates with ridership data for these lines as they are the highest of each respective agency. For New York City Subway, the most customer issues, in order, occur on the No. 1, 2 and 7 trains. Customer Twitter usage may be more influenced by the seasons rather than by ridership during a typical year of service; with higher tweet frequency and term frequency occurring during the end of the summer, and throughout the holiday season.

17. Key Words	18. Distribution Statement No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22161. http://www.ntis.gov		
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 46	22. Price

Form DOT F 1700.7 (8-72) Reproduction of completed page author

Utilizing Social Media Data for Estimating Transit Performance Metrics in A Pre- and Post-Covid-19 World

PI: Camille Kamga
The City College of New York
0000-0002-9223-700X

Sandeep Mudigonda
The City College of New York
0000-0003-1734-673X

Rodrigue Tchamna
The City College of New York
0000-0001-8205-988X

Richard Kish
The City College of New York

DISCLAIMER

ACKNOWLEDGEMENTS

The project team at the City College of New York would like to acknowledge everyone that helped make this project successful. We would like to thank Shri Iyer and Dr. Kaan Ozbay at C2SMART for the opportunity provided to successfully perform this project.

Executive Summary

This study explores the benefits of text and sentiment analysis of Twitter data, as well as the applications of such analysis towards understanding customer sentiment, transit performance, and the relationship between the two. Ultimately the objective of this research is to determine which performance metrics customers most frequently mention and are most passionate about, and where these sentiments are directed when riding with Metropolitan Transportation Authority New York City Transit and how these compare and contrast to the performance indicators reported by the agency. This is carried out by analyzing a year's worth of Twitter data collected from users mentioning their transit experience and recognizing trends regarding sentiment. These trends are then compared with the data reported by MTA in order to find a correlation. The results show that customers mostly tweet negative experiences over positive experiences found with service and are most frequently critical about wait time and delays. These are important factors to consider in a metro as unique as New York's. The data analyzed also gives insight on where the most customer issues occur; most notably the Hudson Line when riding Metro-North Railroad, and the Port Washington Line when riding Long Island Rail Road. This correlates with ridership data for these lines as they are the highest of each respective agency. For New York City Subway, the most customer issues, in order, occur on the No. 1, 2 and 7 trains. Customer Twitter usage may be more influenced by the seasons rather than by ridership during a typical year of service; with higher tweet frequency and term frequency occurring during the end of the summer, and throughout the holiday season.

Table of Contents

List of Figures

Section 1 Introduction

It is commonly accepted that public transport provides an essential service for sustainable future and many cities across the world have realized that a transport system reliant mostly on personalized modes is not sustainable. New York City (NYC) has a highly multi-modal urban transportation system. Various transportation services such as fixed route public transit systems, on demand taxi and personal bike sharing are available. However, service and availability of modes differs between different regions and neighborhoods. Among these modes of transport, the transit system plays an important role as it is the large system which provide mobility in a sustainable manner for residents, commuters, and visitors in the city. Therefore, it is very important for city's administrators to provide transit services that continue to attract riders and consequently remove some vehicles off the roadways. Transit operators must develop some key metrics to measure the performance of the transit service. One of the principal metrics are transit users' perception of the provided services and understanding customers sentiments to-wards the transit service could help capturing the user's perceptions.

The goal of this study is to propose a methodology using sentiment analysis to explore and extract mobility patterns using both openly-available mobility data in New York City and user's posts on social media. Using data from several modes of travel in NYC such as public transit (subway), bicycle-sharing system, and taxi, travel patterns of different demographic characteristics will be analyzed. Several factors that affect these travel patterns will be inferred. Appropriate planning and operational aspects for enhancing mobility will be proposed.

To ensure service is kept up to par with the expectations of a transit agency, extensive information pertinent to the quality of service is collected and tracked throughout its transit systems. With ever-growing accessibility to the Internet, it is easier than ever for transit agencies from cities across the world to be more transparent with their customers about service performance by releasing their datasets to the public and researchers. As an example, in New York City, the current dashboard created by the Metropolitan Transportation Authority (MTA) allows for better communication to customers on the status of service. Most of this information is provided using MetroCard entry data (MetroCard is the fare card used for transit in New York City). The data is available for train capacity, wait times, travel times, and station environment among many other metrics [22, 26]. While transit agencies set their own expectations about the performance of their system, the users have their own series of expectations and opinions when riding the transit system. These expectations have been collected by surveys and more recently and on a greater scale, on social media. Surveys and social media allow customers to express their own sentiment regarding service and allows insight on factors that satisfy customers of transit.

Surveys have always been a useful and conventional way of collecting customer experiences and opinions, especially those involving feedbacks from their transit experience. However, given the requirement of voluntary participation and the time it takes to do surveys, transit surveys generate scantly turnout and usually do not accurately represent the target customer population. Therefore, it is important to seek other sources that could generate customer sentiment on a scale closer to that of the area's actual population. With over 145 million worldwide daily users, Twitter has proven to be an easy-to-use and accessible social media option where users can instantly upload their thoughts and opinions to the Internet from anywhere in the world. Public transit is an integral part of everyday life for hundreds of millions of people. Public transit operators could partner with Twitter Inc. to utilize their expansive user base data and create metropolitan-scale potential for sentiment analysis of customer's perceptions. This study explores the benefits of text and sentiment analysis of Twitter data, as well as the ap-plications of such analysis towards understanding customer sentiment, transit performance, and the relationship between the two.

To accomplish the objectives of this study, this paper is structured in five sections. Following this introduction, the authors performed a literature review on sentiment analysis of transit users to understand the state of the research in this area and results from prior studies on user's perceptions regarding transit services. The review of literature is followed by sections on the methodology and the results of the analyses performed. Lastly, a section on key findings of the study concludes the article.

Section 2 Literature Review

Understanding the key findings of studies on both Twitter data and surveys is important to planning and conducting customer sentiment analysis using transit-related tweets. A handful of techniques are used across various studies and fields in order to define transit riders' experiences through customer feedback and social media use.

Text analysis is a key component of analyzing Twitter data. A Twitter data mining study (6) on transit in the Toronto metropolitan area used a transit dictionary and thesaurus to find relevant terms to the topic at hand such as travel time or safety. The top factors that this study found important to customers in order are location of service, timeliness, information, seating availability, travel time, personnel/quality of service, safety and security, and vehicle maintenance/upkeep.

A study on sentiment analysis of LA Metro feedback found that the Blue and Green lines had a relatively negative sentiment from customers, while the LAX, Expo, Purple, Gold, and Red lines had received a neutral to positive sentiment from customers. Specific lines had varying sentiment throughout the week. For instance, the Blue line had positive sentiment on Mondays, while receiving negative sentiment on weekends. Notable keywords carrying sentiment included "delay, disable, dies, fatal, and beating" (10). A similar study utilized the same approach to sentiment analysis using text, while also finding words that may boost or reduce sentiment such as "not, won't, couldn't, shouldn't, wouldn't". Emoticons were also included in sentiment calculation (3). Another study took Twitter usernames of major airlines and specifically included emoticons and a custom sentiment dictionary in their research (13).

Date and time analysis may reveal when the most tweets are generated, indicating lapses in service, increases in ridership, or traffic disruptions for example (22). A Chicago-based study that delves into micro-events found that positive sentiment was highest around 7-8AM and 6PM, while negative sentiment was highest around 8AM and 6PM as well. The proportion of positive sentiment was highest at 4PM and 12AM, with the most tweets generated around rush hour (9).

Generally, sentiment analysis studies, both survey-oriented and social media-oriented, provided very similar results while differing vastly in the context of either urban transit or suburban transit (1). For urban analysis, people generally criticize wait time, cleanliness, safety and comfort, and fare cost. There wasn't much difference in sentiment between the two types of transit; however, both received generally more negative sentiment than positive. Younger generations were less satisfied with transit usage, although they are the majority of customers (11). Customer sentiment of safety on transit is predictable by the sentiment of safety in the customer's home neighborhood (4). People also cared whether public transit served their neighborhood, and what the closest access to transit was from their home (14). In urban environments, when customers generally have a positive transit experience; not only does ridership increase, but so does the general life satisfaction of the rider (5). For suburban transit, the top criteria expected include itinerary accuracy, system safety, cleanliness, passenger comfort, servicing, passenger information in that order (21).

It is important to understand the different factors and history that distinguish and define transit systems across the world from one another; how these factors relate to the transit rider's customer experience can influence and support a region-specific study. MTA has had to face many challenges specific to New York, including deficits and natural disasters (i.e., 2017-18 transit crisis and Hurricane Sandy), and has a vast history of changing customer service to accommodate for newer circumstances (i.e. MTA "Fast Forward" and Sandy Recovery & Resilience Program). Since the system is comprised of the original IRT, BMT, and IND railroads, transfers are often necessary, making timeliness and frequency the highest concern to customers. Because of this, an MTA study conducted research on a "reach and match" algorithm used to observe headways and calculate wait times for their station countdown clocks (12).

Similarly, another New York study focuses on the relationship between additional passenger wait time, and headway timeliness in order to represent the customer perspective of added travel time. This data is collected using MetroCard data from linked trips (subway to subway/bus). The study expresses interest in implementing these metrics during weekend service when service is affected by track maintenance (7).

In India's capital region, Delhi, the uniquely implemented odd-even traffic policy was subject to mixed criticism on the Internet from motorists. Negatively connotated sentiment used words such as: gimmick, failure, dangerous, vile, choking, and jam. Positive sentiment was received as well, with words such as: smooth, support, and good (2).

As social media data is constantly expanding at an increasing rate, the effect that social media has on customer sentiment has also been studied. Studying the psychology of social media may aid in the understanding of how the public perception of customer service may be biased based on trends and what they read on social media. A Facebook study looks into this phenomenon by correlating sentiment with likes and shares on a post. In a study on natural disasters, groupthink is described as a psychological phenomenon used by influential social media users, celebrities or other spokesperson, where they may use their platform to alter the public's perception and sentiment towards a specific issue (18). Similarly, the more likes or shares a post has, the more reputable it becomes. Ex-NYCTA President, Andy Byford, is an example of this phenomenon in New York when he made a positive impact on New York City Transit's image and its customers through his sociable persona, transparency and community outreach during his two-year run.

Raczyncki et al. (2021) studied everyday social media posts and comments on public transit events in Poland. The authors developed a typology of transit information and a transit stop name database in Poland. These data have been used by them to assess sentiment resulting from various public transit incidents.

El-Diraby et al. (2019) used data over a 11-month period from the TransLink (Vancouver transit agency) Twitter account and developed a network analysis of social media users who interact with the account.

compared them to similar data from two other Canadian cities; and, in so doing, contrasted the results in days with disruptions and days with normal operations.

Throughout all these studies, there is clear overlap between factors that affect customer satisfaction. All metrics across the studied transit agencies report similarly differing levels of positive and negative customer sentiment. However, the magnitude of sentiment depends on the type and quality of transportation available in a specific city, neighborhood, or region; as well as other existing infrastructure, and transportation norms associated with the region. An entire study on safety showed that perceptions of safety within one's own home or neighborhood affected their perception of safety when riding transit. Opinions on what transit was useful for, and expectations of customer service also changed depending on the transit agency involved.

Section 3 Methodology

Before analysis could be planned and executed, the necessary data needed to be mined so that the analysis could be engineered around the format of the harvested datasets. First, an application programming interface (API) for Twitter was used to collect a series of tweets, using keywords and tags associated with transit (train, delay, and #MTA, #subway, etc.). The authors used the API for Twitter to collect tweets from the period of June 2015 to May 2016. This collected dataset was made available to the researchers upon be purchased from Twitter Inc. The data file obtained from the API contains extensive information on individual user tweets geocoded for the New York Metropolitan area, including exact time and date, location, username, and the text body of the tweet. These individual specifications for each tweet allow for a variety of possible methods of analysis, as well as easy customization of what can be used and excluded. Second, the MTA performance data was extracted from a New York State government database containing information on measured metrics including the type of metric, how many times a metric is reported in a month, the magnitude of the metric, and the subsidiary agency of the MTA for which the metric applies.

To get the Twitter data into a more manageable format, tidy text tools and analyses within R Studio were used to extrapolate each tweet as a separate entity with further extrapolation of each word from each tweet as a separate entity, and then lastly stemming each word to its base. Each of these stemmed words are summed across the entire dataset then, the term frequencies can be studied. For example, three tweets that use the words "delay", "delays", and "delayed" respectively will yield the word "delay" when stemmed. Therefore, the three tweets will yield a term frequency of three for the word "delay". This narrows down the results closer to one specific metric and makes analysis easier to digest. A "stop_words" function allows for further removal of unnecessary and non-words, such as "the" or "http" and other similar terms. Additional keywords pertaining to news sources and Twitter pages unrelated to customer experience were also removed. Assigning months to the data and grouping it will allow for direct comparison to the MTA performance indicator data.

In order to conduct the sentiment analysis, an ordinary dictionary could not be used. A sentiment specific dictionary was imported to measure user sentiment across tweets. Two prominent dictionaries used in this research were the "BING" and "NRC" dictionaries. While both carry many of the same words, the applications to sentiment analysis varies greatly between the two. The BING dictionary assigns a word in its dataset either a positive or negative sentiment and an attached weight of that sentiment. The BING dictionary provides a straight contrast between the sentiment of words and allows for a clear comparison of negatively and positively charged tweets (Figure 1).

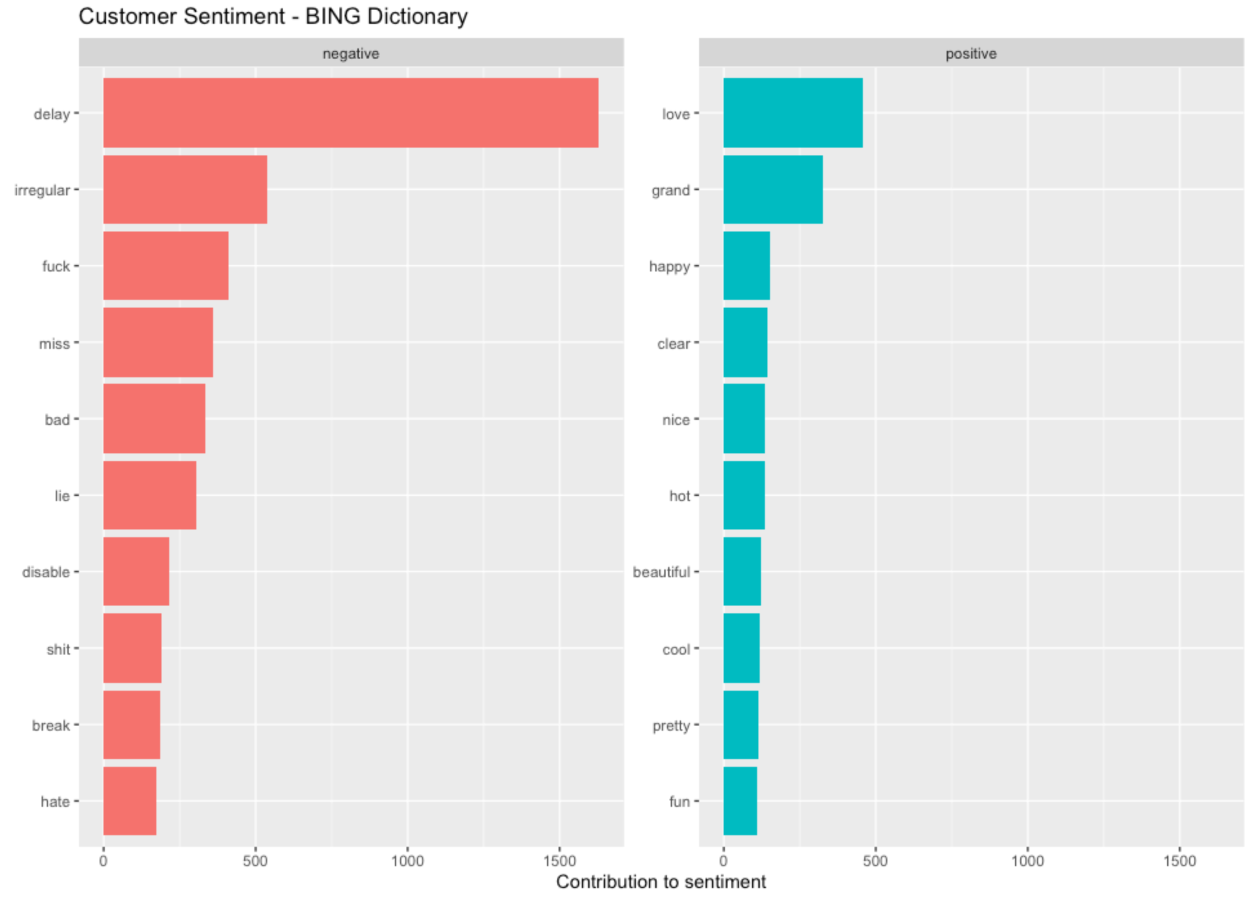

Figure 1 Top 10 words from Bing dictionary and their contribution to +/- customer sentiment

NRC carries less words than BING; however, assigns a word to one, or more, of ten emotions, including anger, anticipation, disgust, fear, joy, negative, positive, sadness, surprise, and trust. Most words overlap with two or more emotions; for instance, "delay" is categorized as anger, anticipation, disgust, negative, and sadness. Although not immediately important to this study, the NRC dictionary provides some valuable insight on the spectrum of emotions Twitter users are experiencing when they are posting about their transit experience (Figure 2).

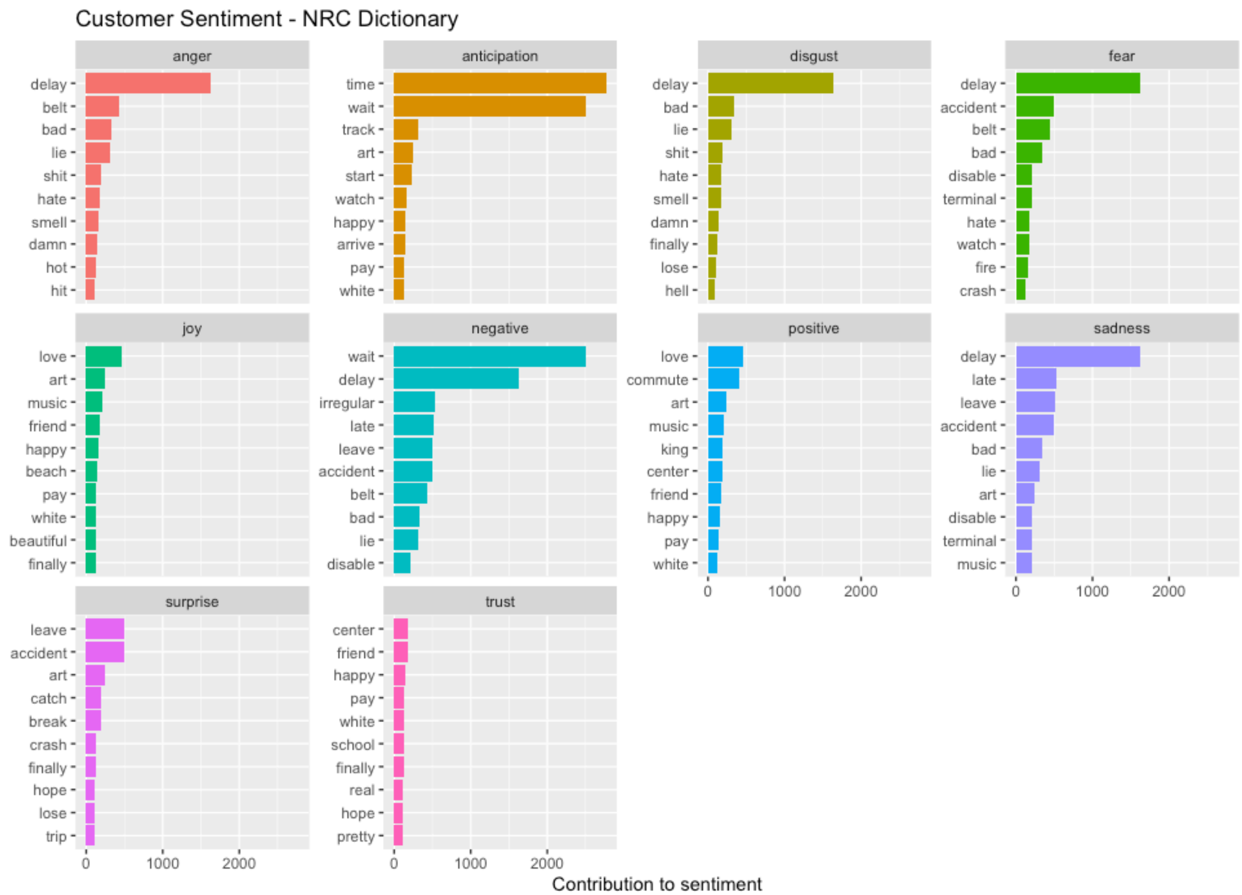

Figure 2 Top 10 words from NRC dictionary and distribution of customer sentiment

The data provided from MTA on the New York State government database holds a similar format to that of the data retrieved from the Twitter API, and therefore is easily organizable and yields informative data. The performance data is only available in month-by-month format, so both the Twitter and MTA analysis will be conducted from a monthly viewpoint. Using grouping, individual metrics can be organized by agency, heavy rail vs commuter rail vs bus, and by individual subway and commuter lines; there are dozens of options when compiling data from the MTA dataset. For example, the percentage of listed wait times met with on time J/Z trains is shown below (Figure 3).

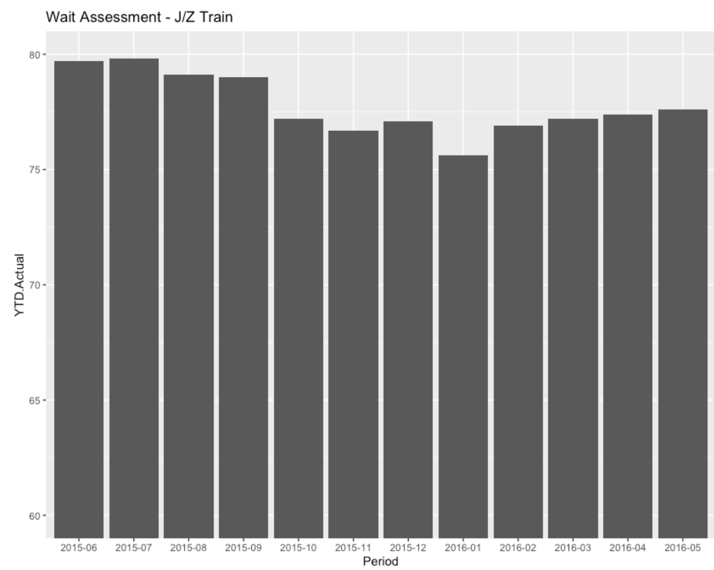

Figure 3 Monthly subway wait assessment- J/Z trains. June 2015-May 2016

Section 4 Results

The results of this study are split into two sections, MTA performance analysis data, and Twitter sentiment analysis data. Both are crucial to this study and are jointly analyzed in the Key Findings section. Through manipulation of both the MTA and Twitter data, plots and figures make use of the processed, filtered, and stemmed data. As illustrated in Figure 4, it is an application of tidy text tools which depicts words organized in a word cloud, with positively weighted words at the bottom, negatively weighted words towards the top, and the term frequency represented by the size of the word itself. Implementation of the customized transit-oriented dictionary creates a different word cloud with all irrelevant words excluded (Figure 5).

Figure 4 Word cloud depicting term frequency and +/- sentiment of transit-related tweets

Figure 5 Word cloud using transit dictionary

Note that a majority of words found in Figure 4 may not be related directly to service, but rather overall experience while riding public transit, including unrelated events. Figure 5 is a visual representation of the words that were selected for the customized transit-oriented dictionary. Words were selected using a term frequency function to count the total amount of times a word was used throughout the Twitter dataset. Only top words related to transit were hand-picked from the function and placed into the separate, customized dictionary. Notice the drastic decrease in non-transit related terms in the new word cloud (Figure 4, 5).

4.1 Twitter Sentiment Analysis During 2015-2016

This portion of analysis represents the nature of the customer opinion and response to their transit experience. It may not accurately reflect the performance provided by MTA, nor the performance indicators reported in the public dataset. It is important to remember that the data used in this analysis is from June 2015 to May 2016 and may also not represent customer experience and response from other years or samples. This data also does not account for periods of time with unpredictable variables that may affect user experience, such as union strikes, inclement weather, natural disasters, or pandemics.

The terms from the customized transit dictionary are grouped by sentiment and each term counted. The greater the term frequency for a word, the more the word contributes to either positive or negative sentiment (Figure 6). The sentiment of the average Twitter user was mostly negative, with a smaller percentage of tweets representing positive sentiment of transit experience (Figure 1, 6). The specific metrics being mentioned can be found through metric-specific terms, such as “wait”, “delay”, and “late” for the timeliness of service. These terms take up a vast majority of the transit-related terms used by Twitter users, which shows that customers mostly tweet about service timeliness. On the positive side, people notice the good things about service, such as system aesthetic, train speed, station cleanliness, and safety in that order.

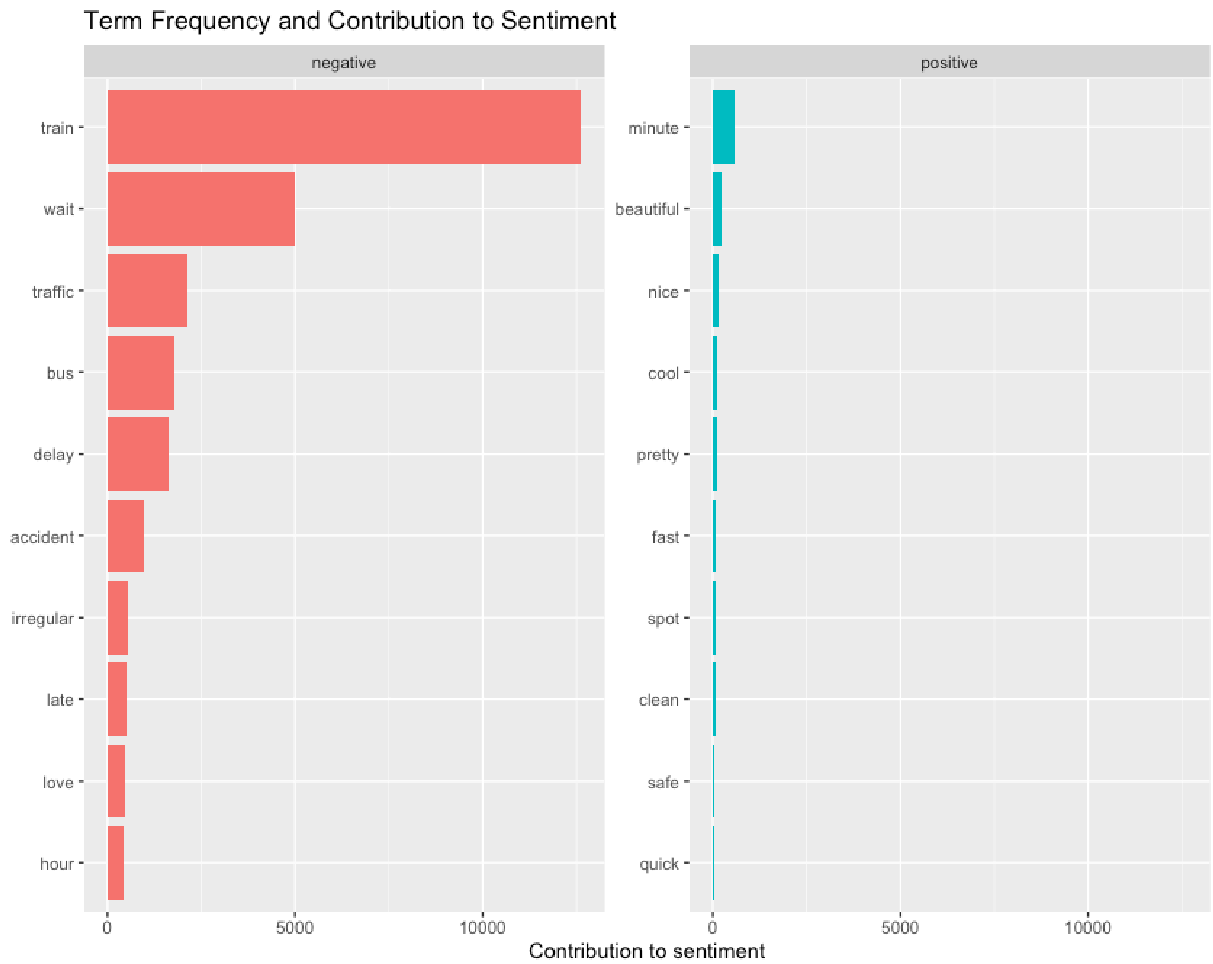

Figure 6 Term frequency of twitter data using a custom transit-oriented dictionary

The issues that Twitter users report on may change throughout the year in response to changes in the seasons and respective changes to service, as well as reasons for using transit. Some words see no difference in term frequency from month to month, such as time-related words. Some words see no instances at all during certain months, like “irregular” which shows up only nine out of the twelve months, and “accident” which shows up only four of the twelve months (Figure 7). Overall, there is no clear change in different metrics taking priority for Twitter users; however, more research into this with more terms may be useful for future research.

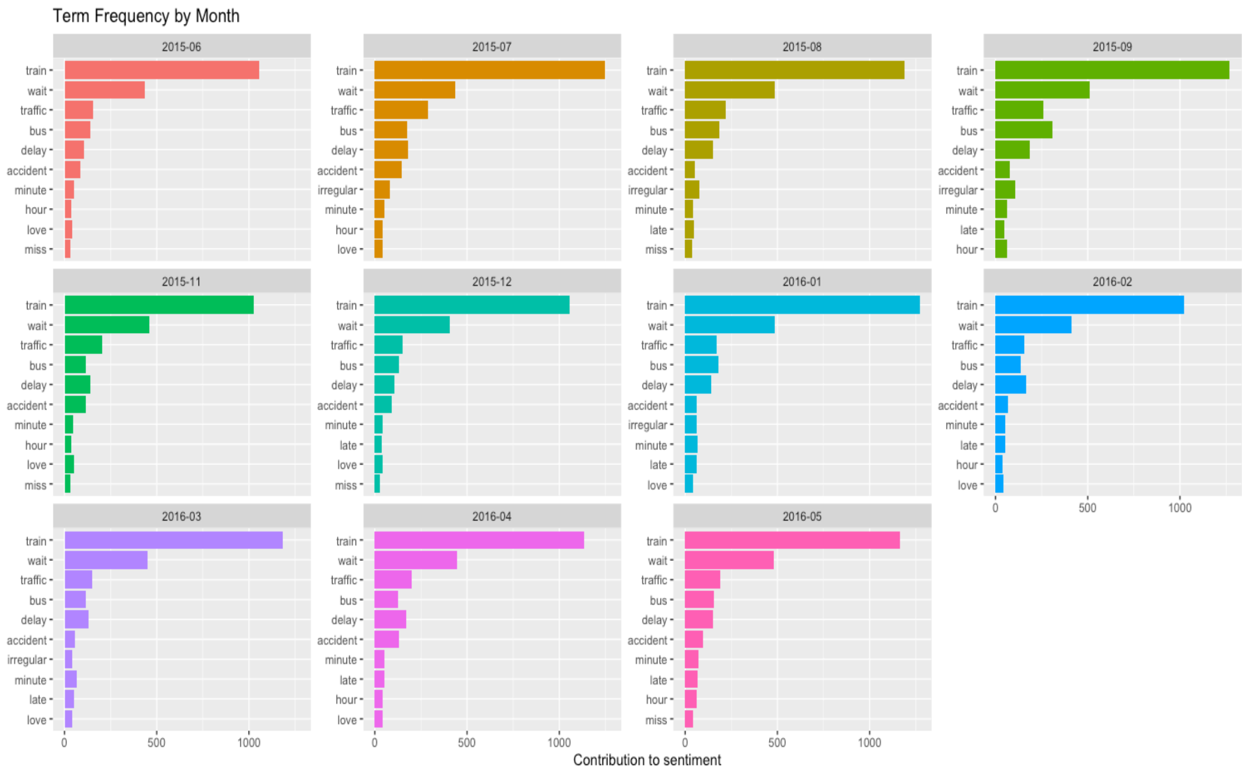

Figure 7 Monthly term frequency of twitter data using transit dictionary

Term frequency can also be grouped by the performance metric each word is related to. Again, the Twitter data contributes the greatest to time-related customer feedback. Following far behind in second are tweets that mention traffic and bus-related incidents. The metrics that are most important to customers based on this data in order are wait times, service quality, crowdedness of trains, safety in stations and trains, appearance, and personnel behavior (Figure 8).

Visualizing the metric distribution of tweets may provide more insight into overall customer Twitter behavior. Though the exact context of each tweet cannot be determined through what normal tidy text analysis is used for, separating the data into metrics can allow for easy organization of tweets when they are able to be sorted by contextualization.

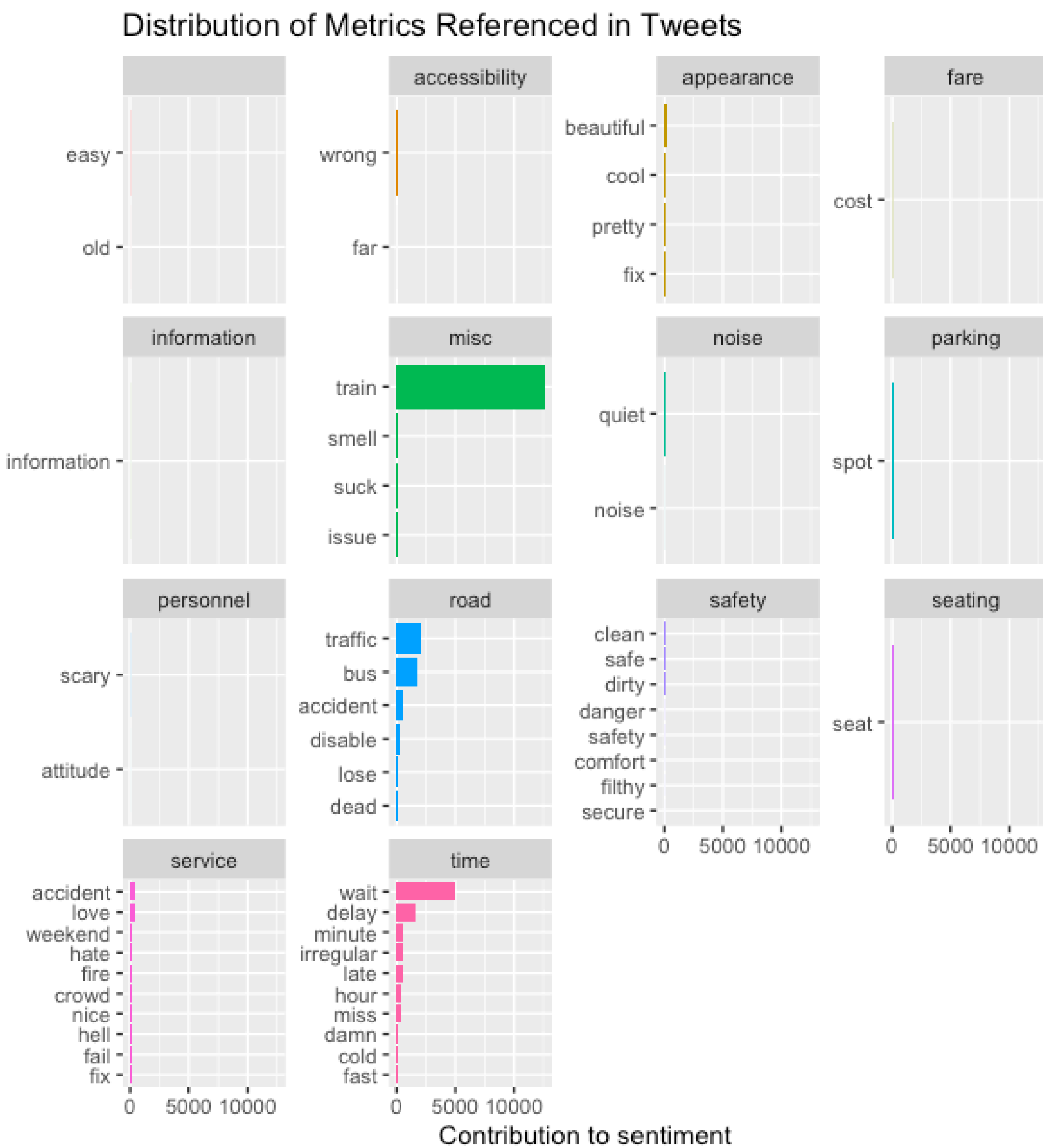

Figure 8 Contribution of term frequency to performance metrics

The usage of each individual term can be tracked by month and compared to other metric-related words. Similarly, to the overall term frequency, words related to service timeliness are used much more frequently than words related to other performance metrics (Figure 9). While the term frequencies jump around a lot throughout the year, the trend seems to follow the same path for all terms. This may be because of numbers in ridership, as well as active Twitter users during each month, may be inconsistent and therefore may not reflect any noticeable periods of time where certain issues are more prevalent to customers.

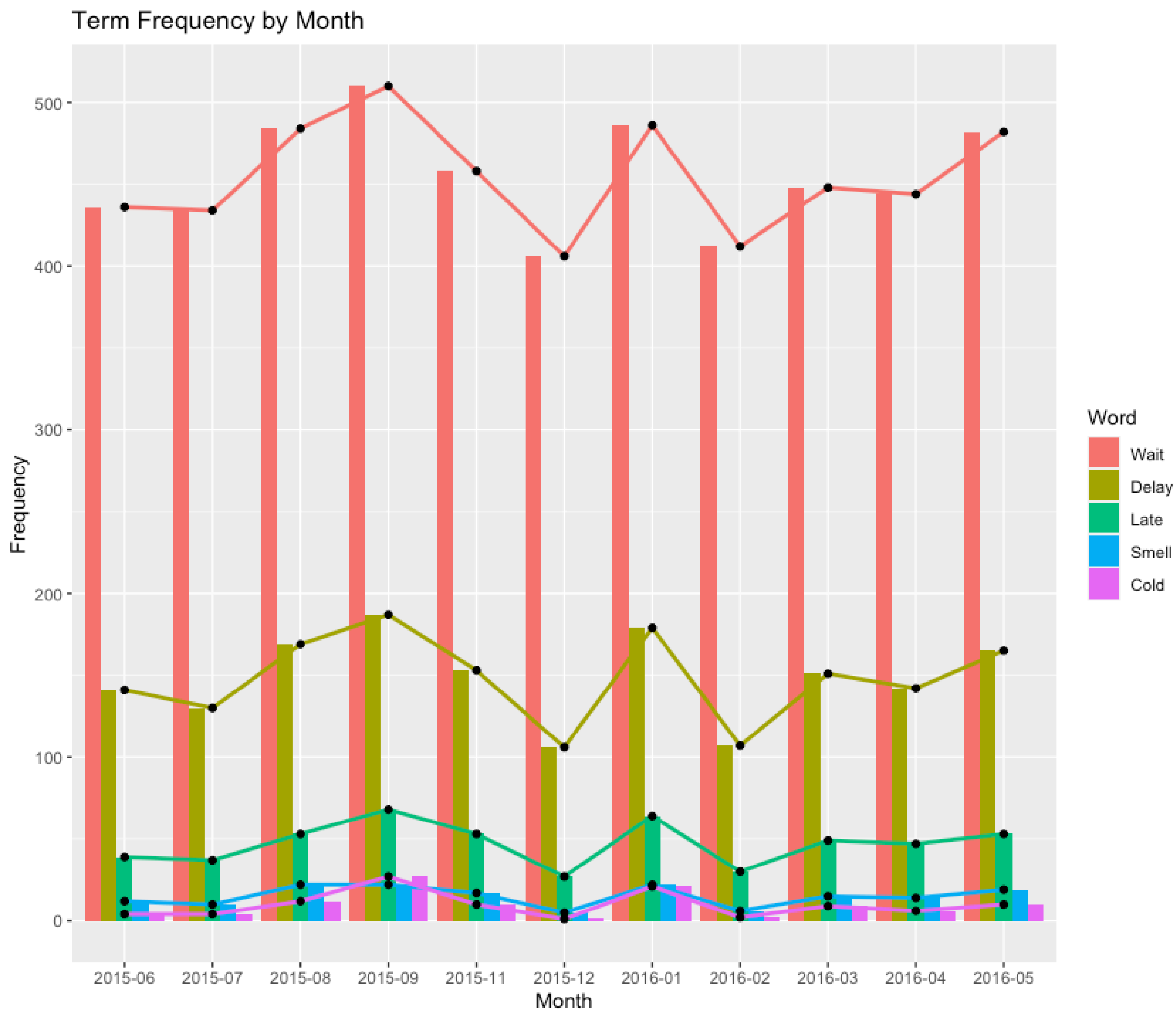

Figure 9 Comparing term frequencies by month for specific metric-related words

Since time-related terms are abundant in comparison to terms related to other metrics, those other words are compared separately (Figure 10). Again, a similar trend is seen across all terms listed, but the issues customers are experiencing can be listed in order from most important to least important. These issues in order pertain to, station environment, crowdedness of trains and stations, cleanliness, train speed, and lastly, fare cost.

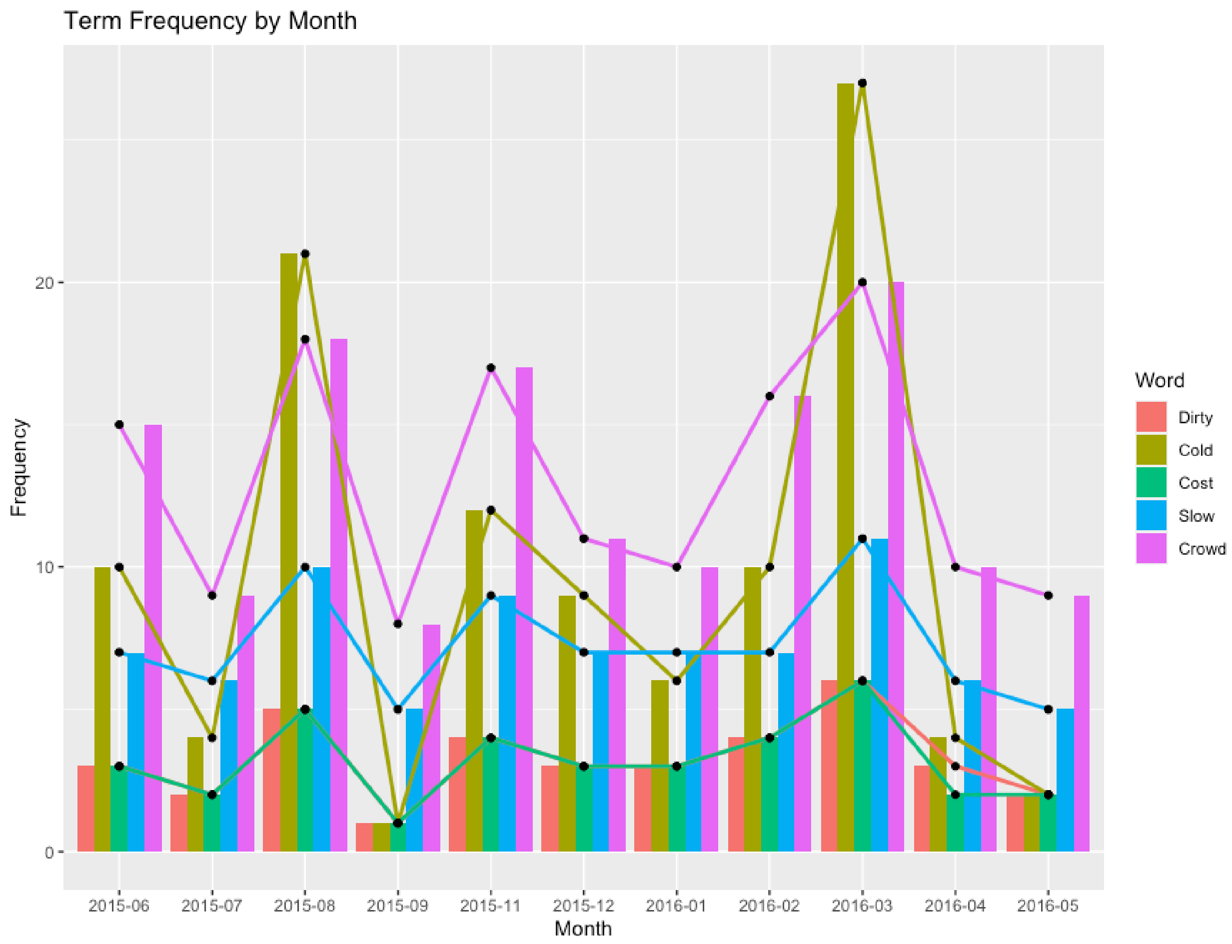

Figure 10 Comparing term frequencies by month for more specific metric-related words

With grouping and the use of a secondary dictionary, the term frequency can be calculated for a specific line or station name, train lines, locations, and any other specification that can be named. For the subway, the term frequency of individual lines can be calculated for the A Division. Because of the way tidy text tools and stemming work, the term frequencies for the B Division of trains, or all the lettered trains, could not be calculated or obtained in any way. In order, the subway lines most frequently tweeted were the No. 1, 2, 7, 3, 4, 6, 5, then Shuttle trains (Figure 11). The tweets carried around 1000 mentions per subway line on average.

The same method of using a secondary dictionary for transit lines worked when figuring out how many times a specific commuter rail line was mentioned in tweets. Only individual words were able to be extracted, so lines with two pronouns in its name like “New Haven”, “Oyster Bay”, and “Pascack Valley” were substituted with “Haven”, “Oyster”, and “Pascack”. The most frequently mentioned commuter rail line by far was Metro-North Railroad’s Hudson Line with nearly 250 mentions. This is followed by the Harlem Line, and Long Island Rail Road’s Port Washington Line (Figure 12). The frequency of these lines were around 100 mentions; around one-tenth the amount of times subway lines are being mentioned.

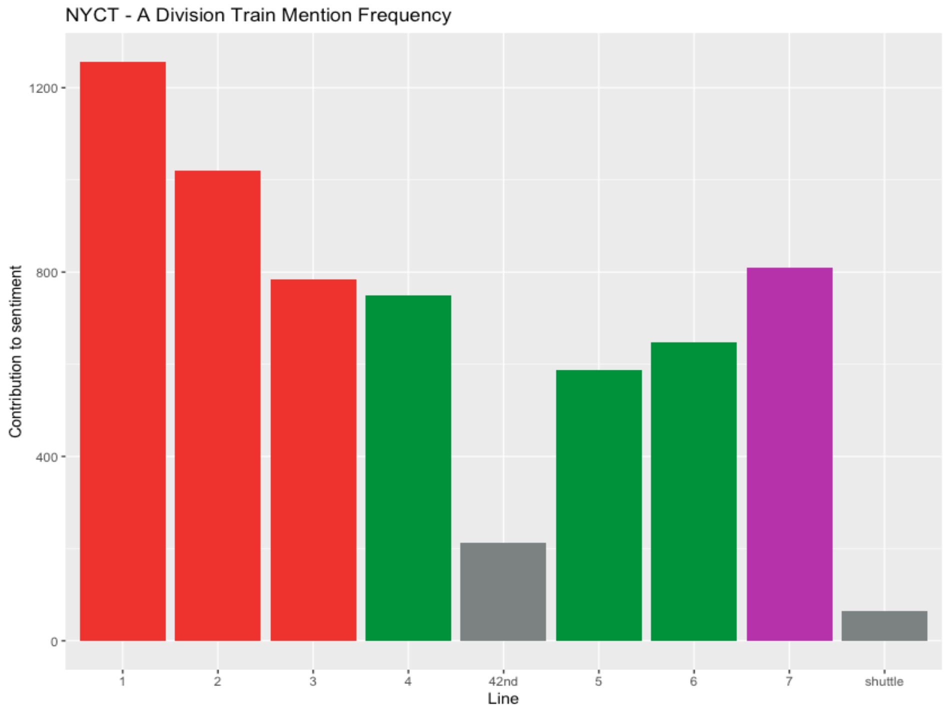

Figure 11 Mention frequency of New York City subway A division trains

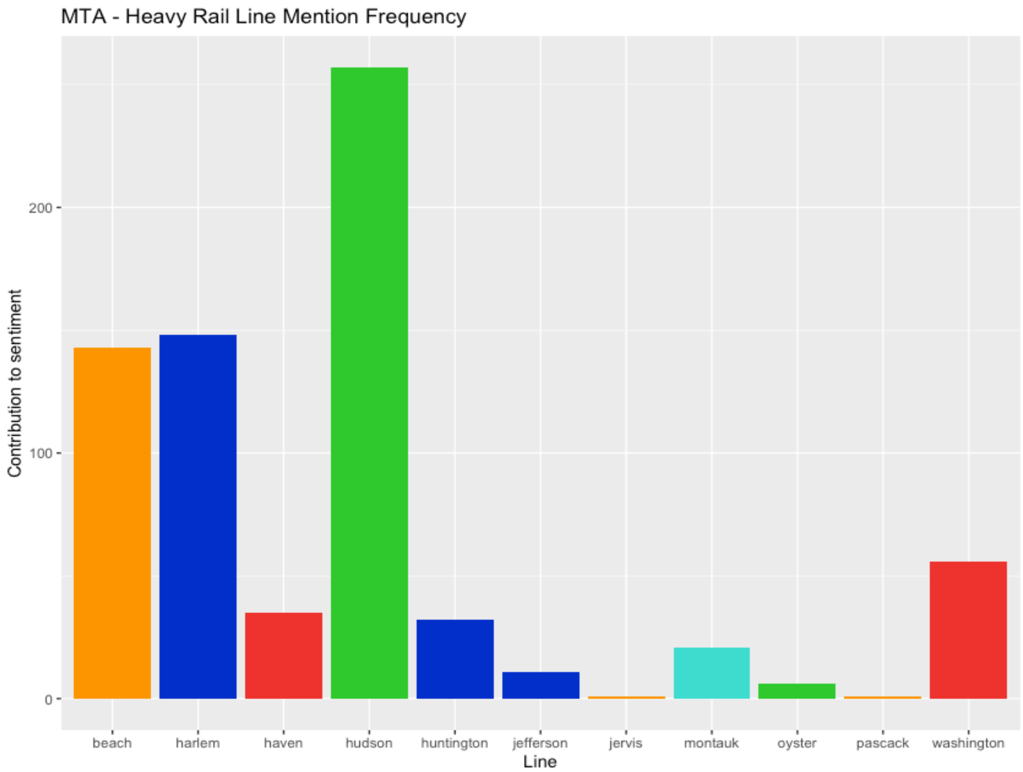

Figure 12 Mention frequency of MTA commuter rail lines

4.2 MTA Performance Analysis

The MTA data is also from June 2015 to May 2016 in order to maintain consistency with the Twitter data and reflection of actual service performance. Performance metric ratings are available for MTA Bridges and Tunnels, New York City Subway, Long Island Rail Road (LIRR), Metro-North Railroad (MNR), and MTA/NYCT Bus. This analysis is cross-examined with the Twitter data analysis in order to find similar or contradicting trends. Since traffic-related tweets were left out of the Twitter analysis, Bridges and Tunnels data was also excluded from the analysis of the MTA data. Also, all rail and bus metrics are generally shared among the agencies.

As discovered in the Twitter analysis, wait times and itinerary accuracy are of the highest importance to MTA customers. Therefore, reviewing the on-time performance of each of the agencies across MTA provides insight to the reasons for customer behavior on social media. The on-time performance is much higher on commuter rail lines than on subway lines (Figure 13). An inverse relationship between customer complaints and quality of service is found when compared to tweet mentions (Figure 11, 12). Below 70% acceptable performance, subway lines were mentioned more on Twitter, compared to over 90% acceptable performance by commuter rail lines, which generated less mentions in the Twitter data.

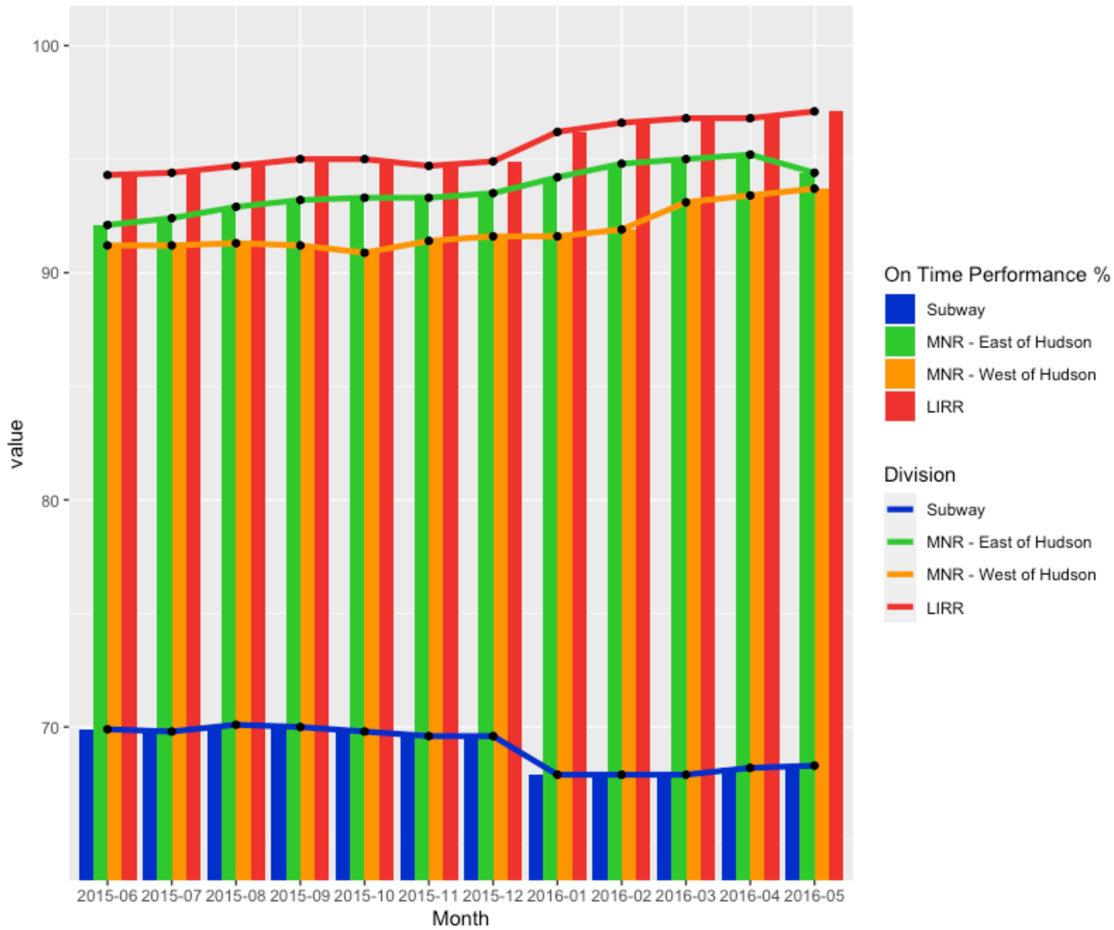

Figure 13 Percentage of on time performances of MTA agencies

Subway wait time assessments allow for comparison of how service differs between individual lines of the New York City Subway. Figure 14 depicts the numbered lines of the subway and the percentage of trains that arrived on schedule for that month. All but the No. 1 train shared a similar trendline path as the year progresses. Throughout the year, the percentage of on-time trains gradually decreases then drops a noticeable amount during the holidays, while the No. 1 train reports higher on-time percentage from January 2016 throughout the first quarter. When compared to the term frequency of the word "wait", "delay", and "late", their highest term frequencies occur during and just after the holiday (Figure 14). Similarly, the wait time assessment sees a minimum in September before slightly increasing before the holidays and is reflected in the increase in term frequency for the same time-related terms for the month of September. Interestingly, the top three train lines mentioned in the Twitter data, are the first, third, and fourth best performing train lines based on wait assessment.

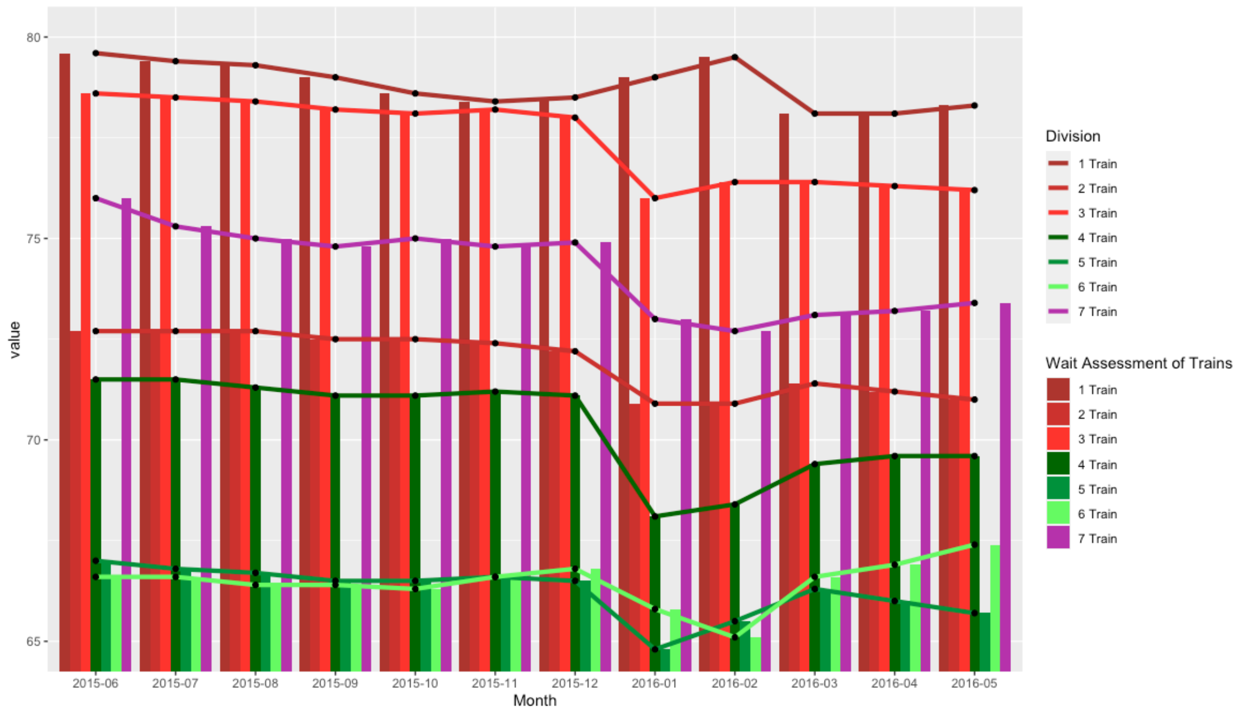

Figure 14 Monthly subway wait time assessment of a division trains

The mean distance between failure of train cars and buses provides insight on the reliance customers can have when riding with MTA and how the quality of their fleet fluctuates throughout the year. As the year progresses, the condition of the fleet improves for bus and commuter rail but decreases for subway (Figure 15). These increases and decreases in fleet condition are not reflected in the Twitter data, which suggests there may be more reliable of a metric, such as ridership, for determining how often metric-specific issues are encountered and reported during each month.

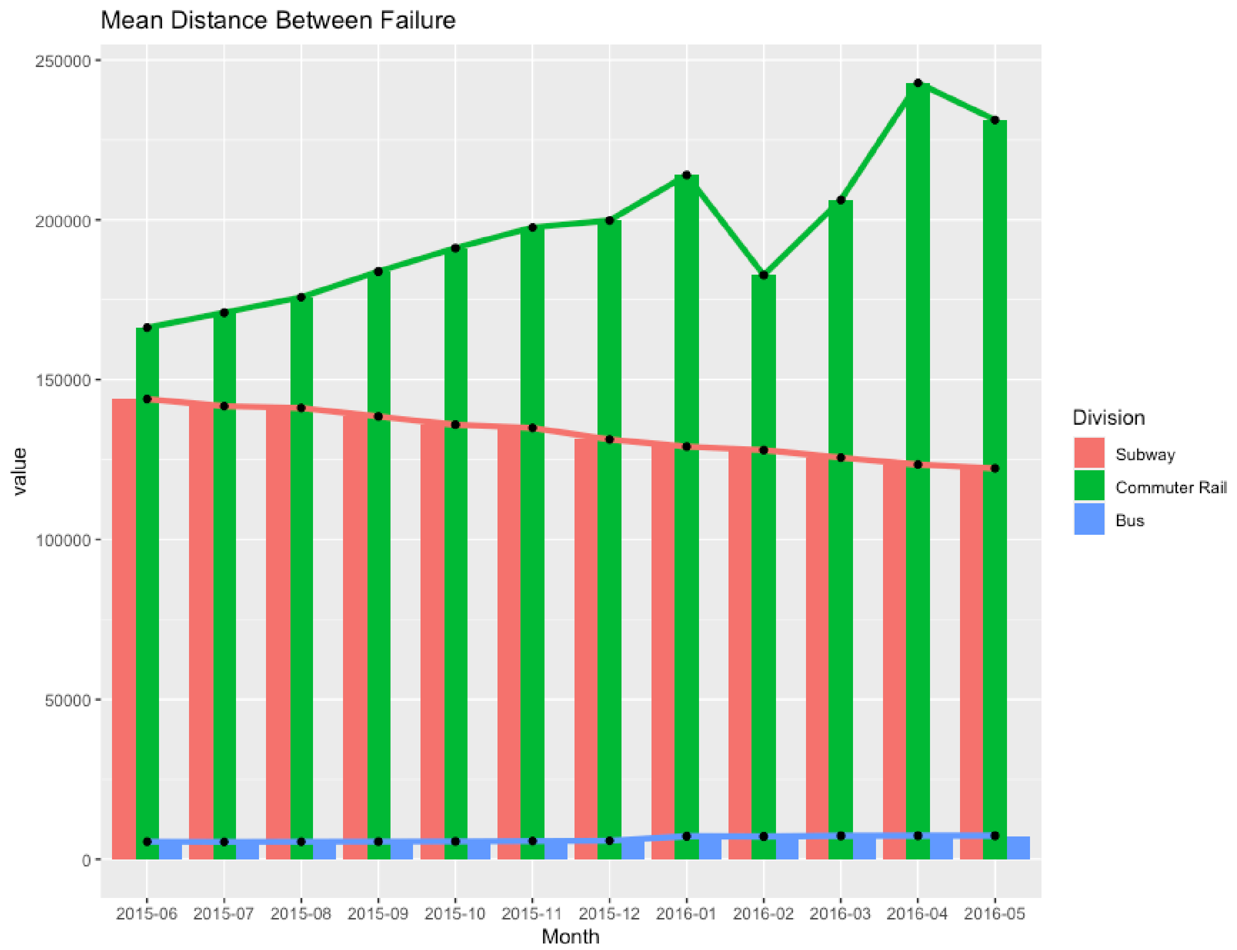

Figure 15 Mean distances between failure of MTA modes of transit

Although safety isn't a highly criticized MTA performance metric, it is equally as important as the accuracy of the itinerary that public transit is safe. Assuring customers that the safety practices are in place may encourage people to choose public transit in the future (4). As the year progresses the injury rate stays about the same across all agencies, until the holidays when a jump is seen in injury rate and eases into a decline throughout the year. Interestingly, the decline starts in different months across all the MTA agencies. Metro-North Railroad sees its decline in injury as early as December into January. New York City Subway and MTA/NYCT Bus sees a decline from January to February. Long Island Rail Road doesn't get its injury decline until as late as March, shortly making a huge dip in the injury rate before jumping back up in April (Figure 16). Further research into what caused these injuries, how they can be correlated with the month they occurred, and if these issues are ones mentioned by Twitter users, can help to incorporate a broader application of sentiment analysis to improving system safety.

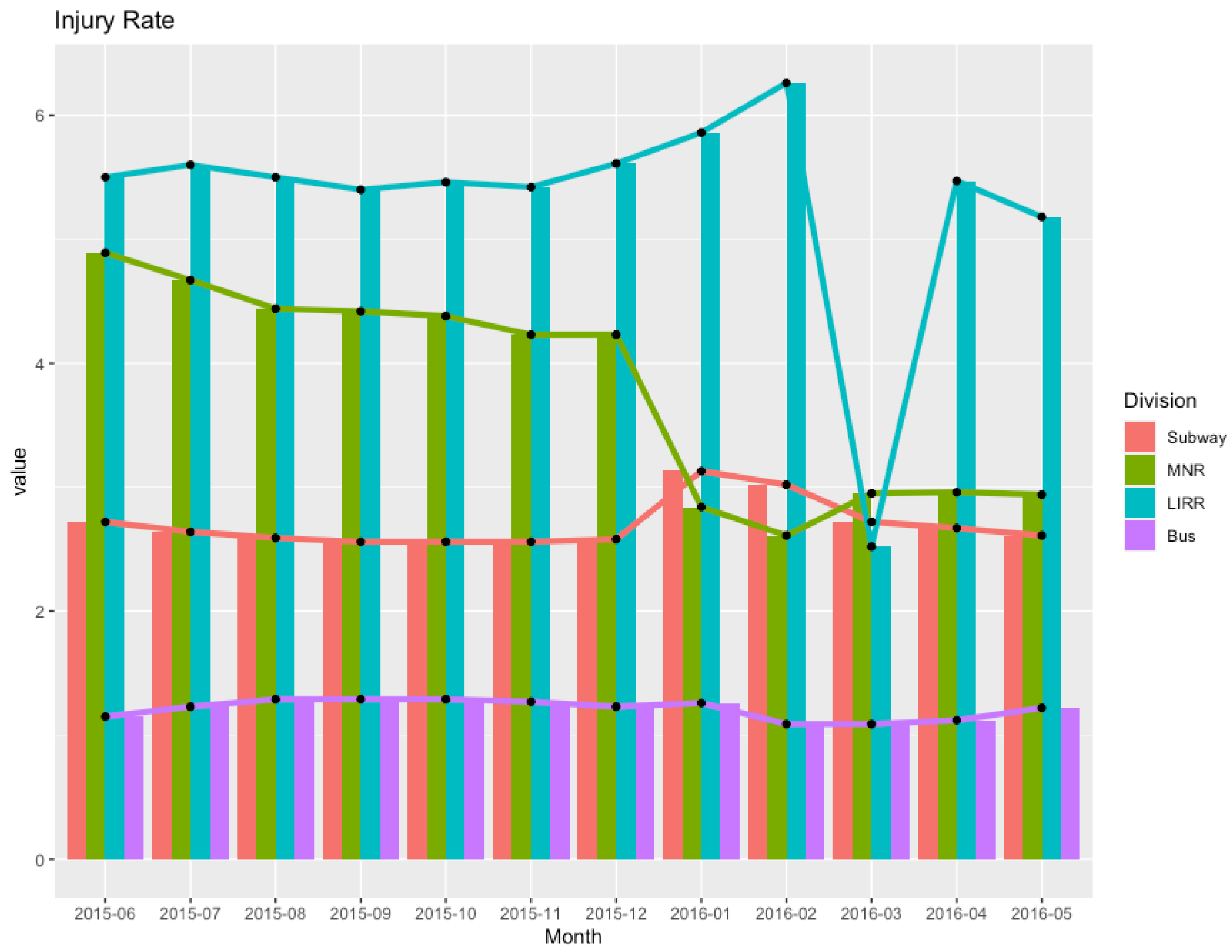

Figure 16 Injury rate between MTA agencies

Ridership data illustrates the crests and troughs of service usage over the year and provides a deeper understanding on how other metrics perform throughout the months. Generally, all agencies saw similar rises and falls in ridership over time, indicating that ridership is dependent on the time of year. Ridership is seen slowly decreasing as the summer ends before increasing during the fall, likely due to the end of summer break and an increase in activity. Ridership seems to decrease over the winter, with a small increase during the month of December. As the days warm up going into spring and summer, ridership numbers see a steady increase. A clear difference in ridership numbers across the different agencies shows which services may require the most maintenance and receive the most tweets. Subway ridership reached an average of 150 million monthly riders and bus ridership reached about 50 million per month. Metro-North Railroad and Long Island Rail Road ridership numbers each averaged just under 10 million monthly riders (Figure 17). Subway ridership numbers are ten times as much as commuter rail numbers, which directly correlates to the number of Twitter mentions between subway and commuter rail users (Figure 11, 12). Ridership appears to have no influence on injury rate, or fleet performance; however, wait time assessments tend to be better during months of lower ridership, with worse wait time assessments during months of high ridership (Figure 14).

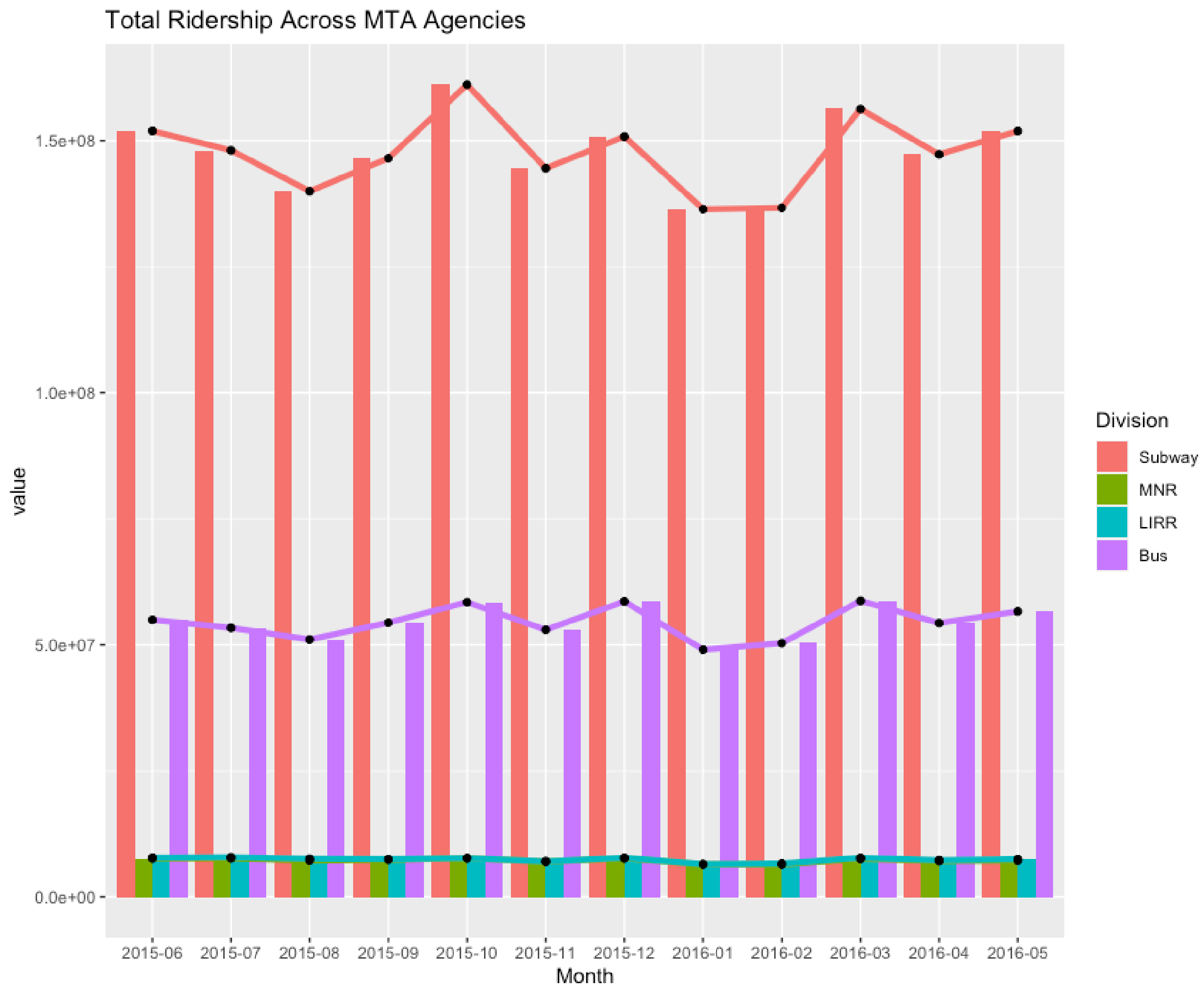

Figure 17 Total ridership between MTA agencies

With a better understanding of the changes in ridership throughout an average year and their effects on metric performance, this same analysis can be conducted in conjunction with term frequencies of metric-specific words. This analysis provides insight into customer issue and feedback trends that may not be accounted for in the performance indicator data. Comparing a subway ridership trendline with the term frequencies of the most common metric-based words used shows that term usage isn't directly linked to ridership (Figure 18, 19). This can be seen during months where increase in ridership is associated with a decrease in term frequency during the months of June-July 2015, as well as an increase in term frequency during the months of August-September 2015. The broadest range this direct relationship of increasing term frequency with ridership is seen is during the months of January-April 2016 (Figure 19). This leads to a suggestion that term frequency numbers may not have to do with ridership numbers, but instead may have to do with the time of year customers of reporting their transit experience on Twitter.

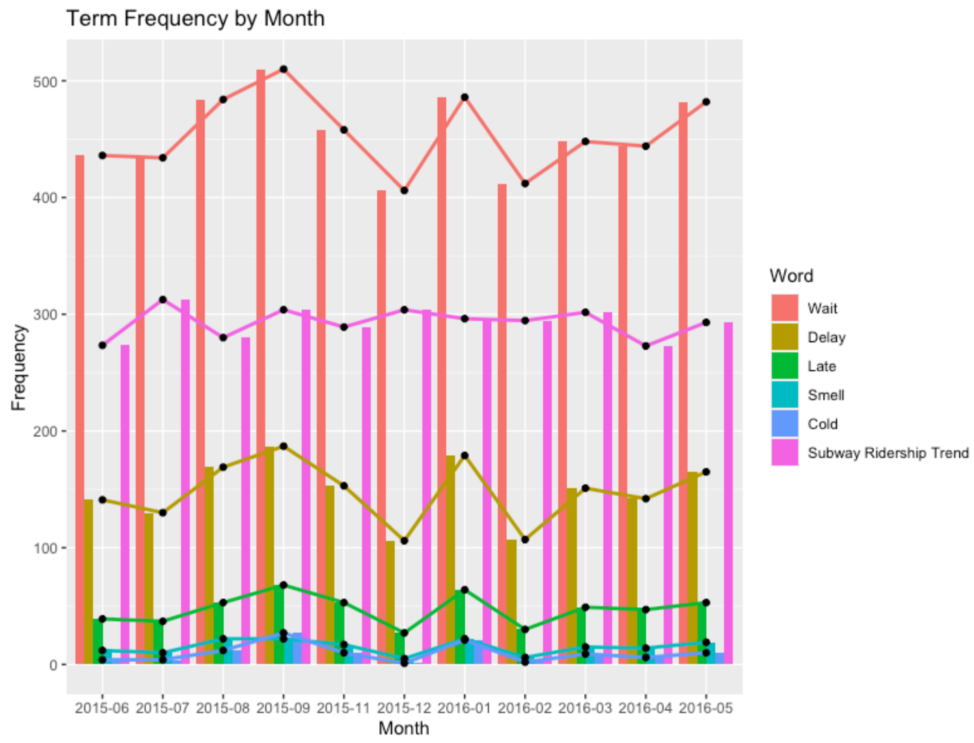

Figure 18 Comparing monthly term frequencies to monthly subway ridership trends

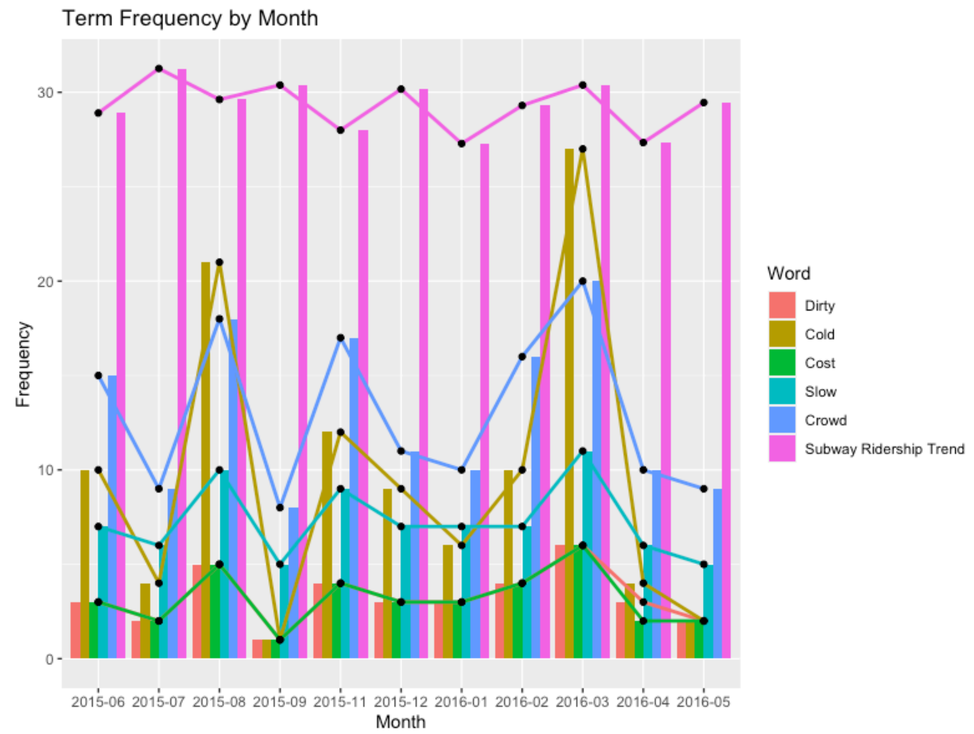

Figure 19 Comparing more monthly term frequencies to monthly subway ridership trends

By normalizing the term frequency with the ridership data, the variations in customer Twitter usage may be visualized. Term frequencies are calculated per million subway riders for each month. Now, the certain customer issues are reflected as a function of the time of year, rather than as a function of ridership (Figure 20, 21). For time-related words, the term frequencies appear to rise throughout the fall, drop as they approach the holidays, and rise again after the holidays. This indicates that more people may be actively using Twitter upon the end of summer break, and again once riders return after the holidays. Perhaps the busyness of New York as the year comes to a close may leave riders with less desire to post about their transit experience to Twitter, whether it is due to crowds or decreasing temperatures. For station environment-related words, there is a direct correlation to the time of year in which these issues may occur. Tweets regarding smelliness of stations and trains have a higher frequency during warmer months, where perception of smell is greater than in colder months. Similarly, tweets related to the cold become more frequent from the beginning of fall to the end of the year and decrease from the start of the new year. Tweets related to crowdedness tend to be higher during the summer months, and during the holiday months. Tweets related to train speed see an increase during the holiday months, where inclement weather may affect service during this time. The frequency of tweets related to fare cost, and visual cleanliness of stations and trains remain relatively constant throughout the year, where these issues may be independent of the time of year.

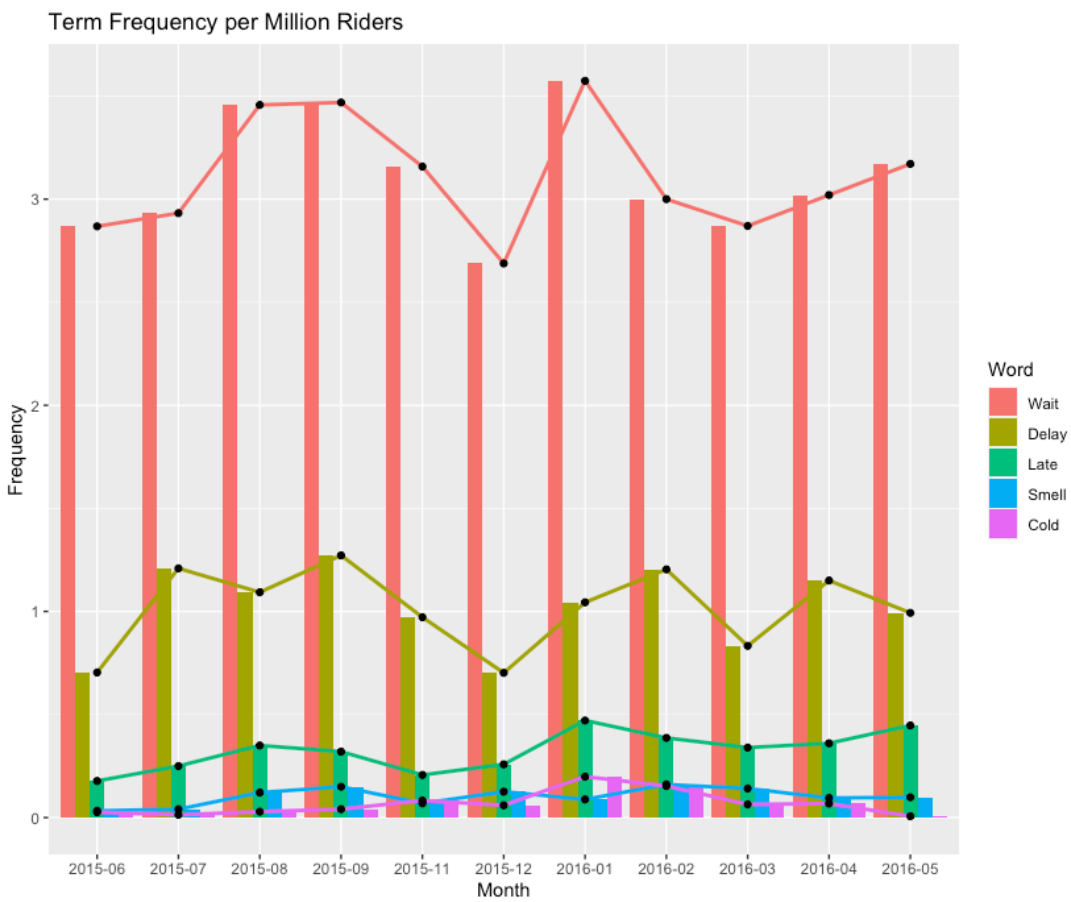

Figure 20 Comparing term frequencies per million monthly subway riders

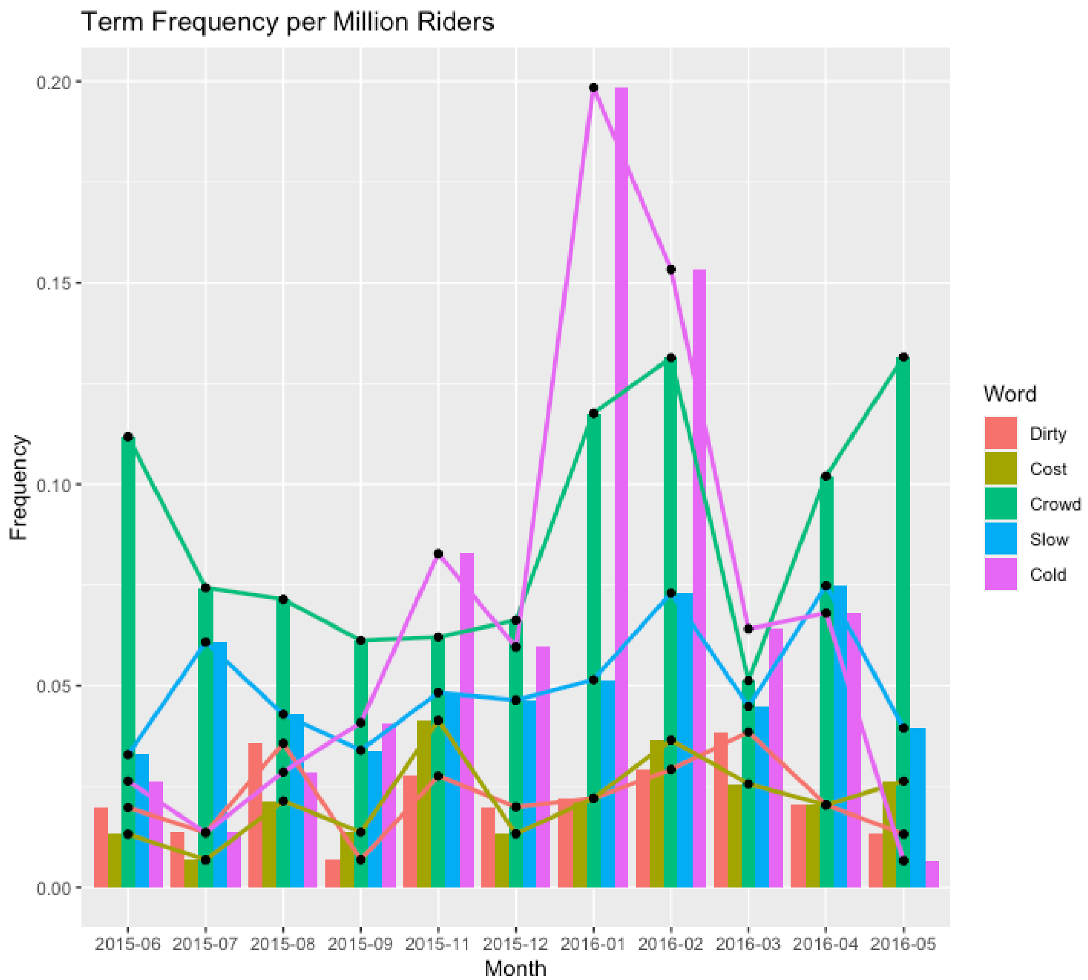

Figure 21 Comparing more term frequencies per million monthly subway riders

4.3 Transit Sentiment Analysis Before/During Covid-19

The data in this portion of the analysis is longitudinally taken from both the 2015-2016 period previously studied, as well as from April 2020 – December 2020, during the COVID-19 pandemic. Sentiment analysis and tweet frequencies from both time periods are studied in order to find differences in trends between the two time periods. The contributing factor to this study is the introduction of an international pandemic and its effects on public transit customers' social media behavior. Since COVID data is taken from April-December, the 2015-2016 data has been cleaned to represent this same time interval. All sentiment and tweet frequency data has been aggregated based on time of day, and time of the year in order to get a broad understanding of how the dynamics of customer behavior change based on time.

Before getting into the sentiment analysis is important to understand exactly how COVID-19 shaped the most important of customer concerns. As the COVID-19 pandemic unfolded, there was much uncertainty with how the virus spread, and how measures were going to be put in place to avoid transmission of the virus as much as possible. Cleanliness procedures were put in place all across the country, with New York City creating a mask mandate almost immediately to reduce COVID transmission. By studying individual term frequencies of tweets from customers during 2015, and the COVID-19 pandemic, it is evident that some customer concerns maintained their rank, while other concerns ranked up or down depending on their relevance to the pandemic. For instance, looking at both 2015 and during the pandemic, wait times were still the number one concern of public transit customers (Figures 22 and 23). Irregular train service, crowdedness, and station environment were all outranked by cleanliness of trains and stations after the introduction of the COVID pandemic, as more customers tweeted their frustrations about the safety of riding public transit during the pandemic.

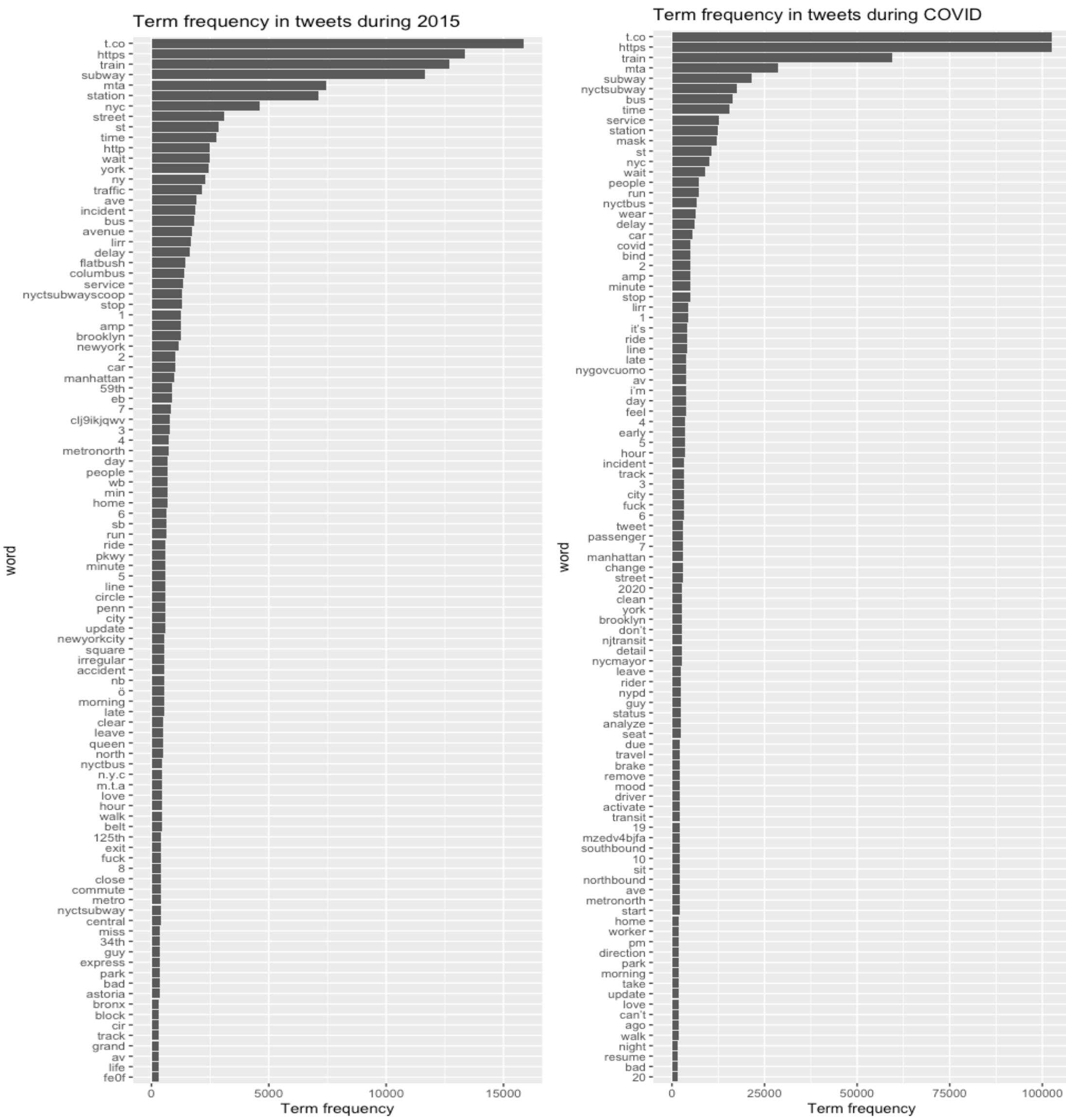

Figure 22 and Figure 23 Comparing term frequency of transit customers and identifying top customer concerns (2015-16 vs COVID)

Additionally, it is important to take a look at how tweet frequency is fluid throughout the course of the day, and the year. First, we will focus on weekdays, as both public transit and work schedules are affected by the COVID pandemic most during these five days. Customers show a dramatic change in tweet frequency between pre-COVID and COVID, especially during the mid-day hours. This is reflected with an increase in tweets during the mid-day during COVID, while there is a decrease in tweets during the mid-day before COVID (Figure 24). This could be due to the increase in hybrid/remote work leaving more time for people to freely tweet during the work-day. Additionally, the near 90% drop in public transit ridership during this time may also contribute customers feeling safer or more comfortable taking out their smart devices and tweeting while riding on public transit. Looking ahead to how tweet frequency differs throughout the year between the two time periods during weekdays, both periods follow very similar trends; however, it appears that tweets related to transit were more common prior to COVID as opposed to during COVID (Figure 25). This makes a lot of sense given that, while Twitter users grew in number from 2015 to 2020, transit ridership was still on a great decline during this time period.

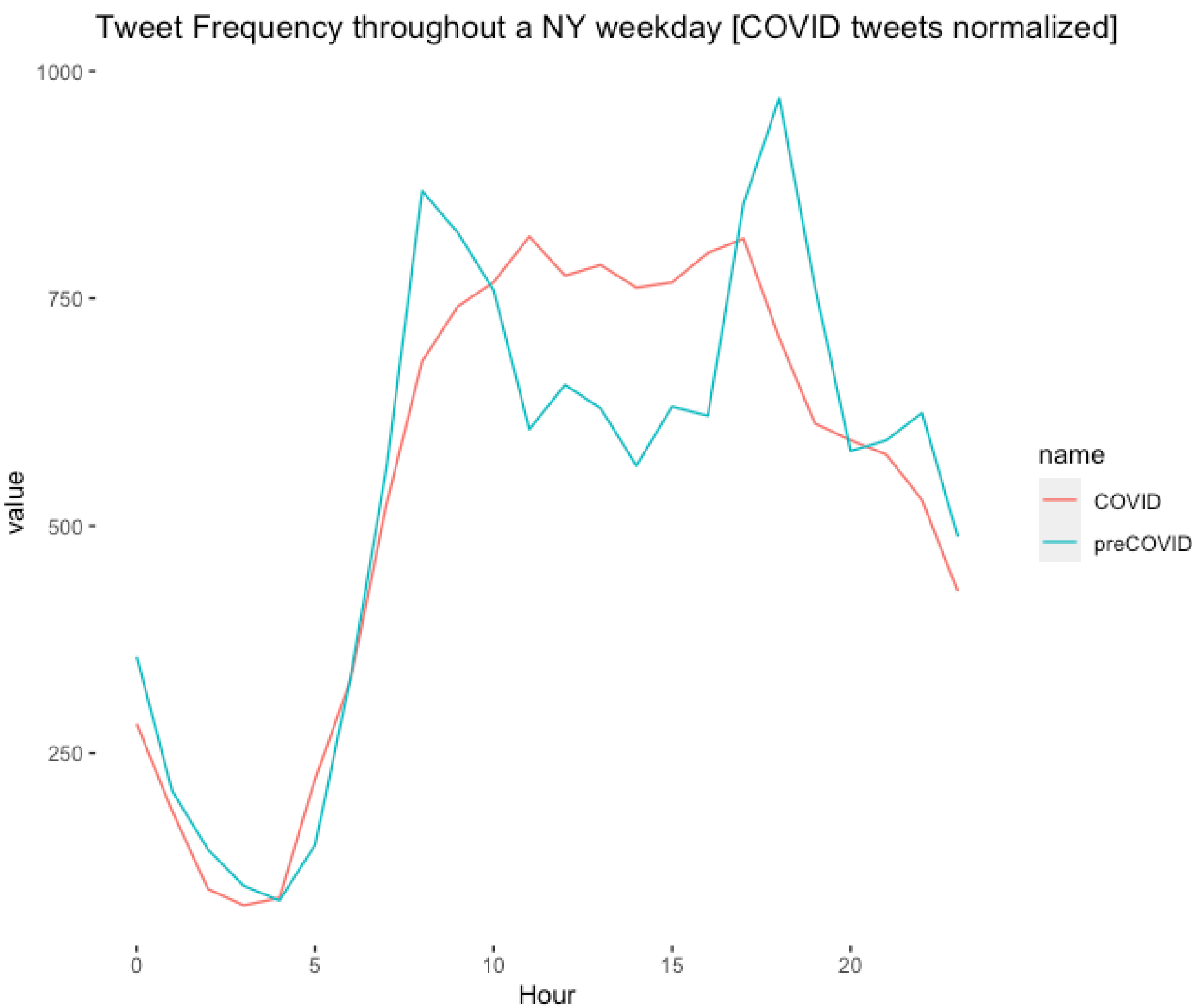

Figure 24 Comparing aggregated tweet frequency during a NY weekday (2015-16 vs COVID)

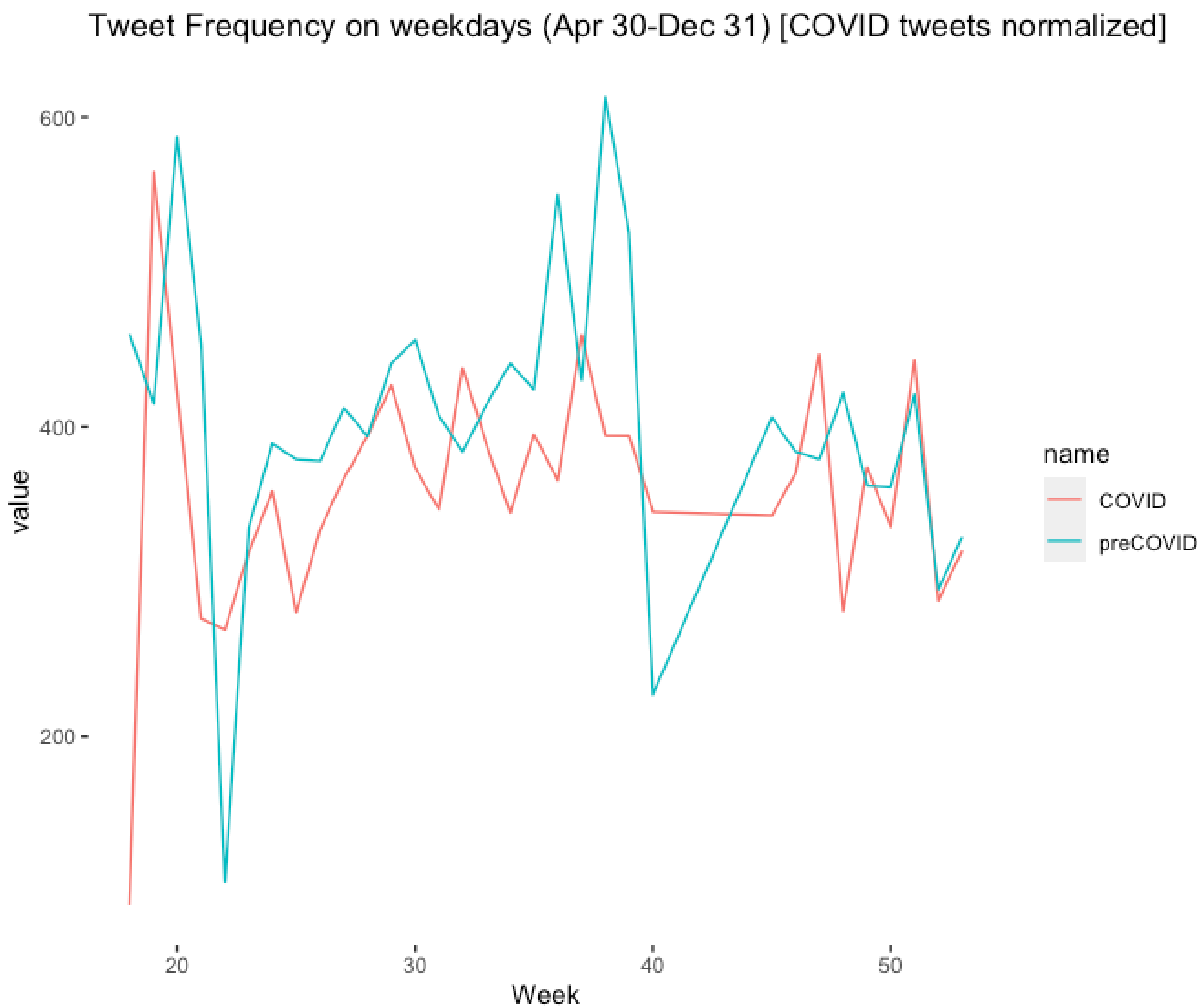

Figure 25 Comparing aggregated tweet frequency during a NY weekday (2015-16 vs COVID)

Looking at weekend data for tweet frequencies, it appears that the COVID pandemic has less of an impact on customer's willingness to tweet about their public transit experiences throughout the day. In fact, the tweet frequency per million for each hour of the day is almost identical from pre-COVID and during COVID (Figure 26). On the other hand, the total tweet frequency throughout the year was drastically affected by COVID, especially during the first few months of the pandemic. While in 2015, tweet frequencies are at an all-time high during the warmer months (starting with April), there is a downward curve in tweet frequencies during COVID, as ridership plummeted almost immediately after the announcement of the public health emergency (Figure 27). As New York City got through the first major wave of COVID and public transit ridership slowly increases back to normal rates, the tweet frequency starts to normalize itself along the same trendlines as pre-COVID measurements, with another peak towards the end of the summer season.

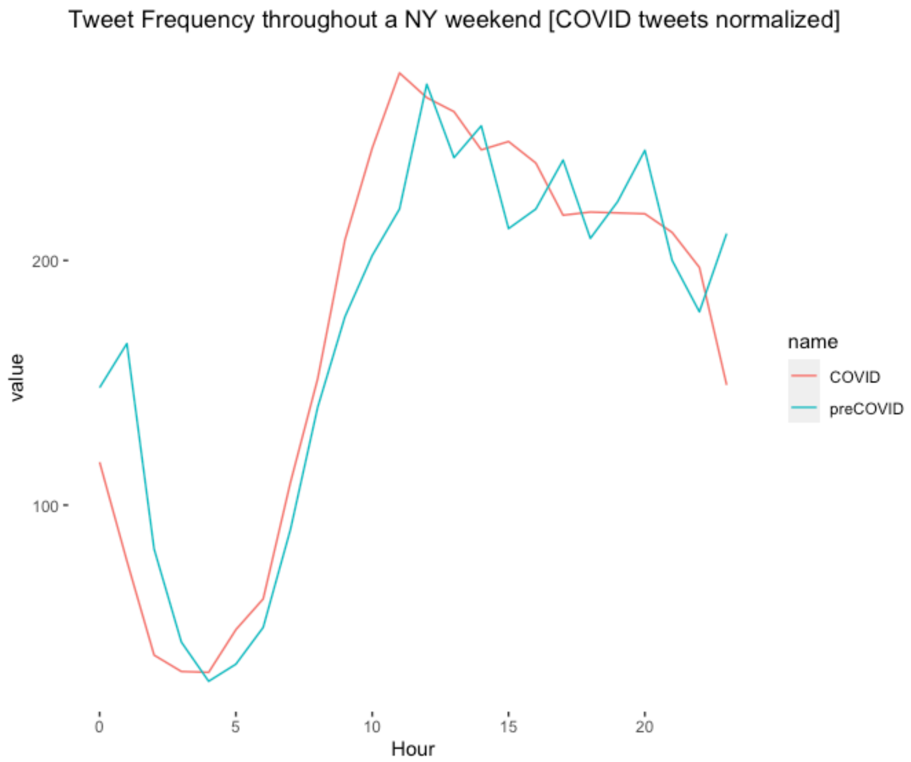

Figure 26 Comparing aggregated tweet frequency during a NY weekend (2015-16 vs COVID)

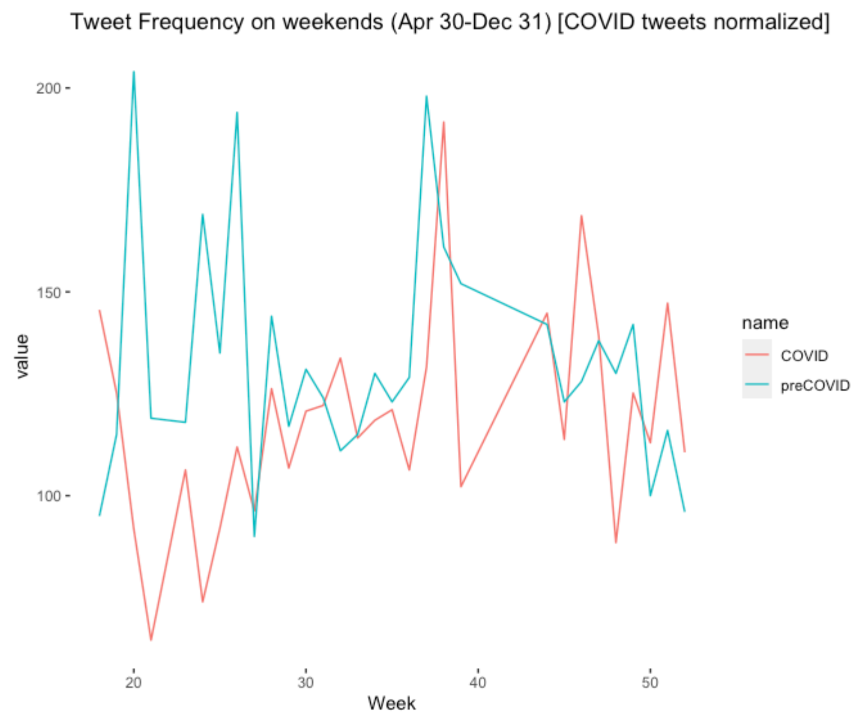

Figure 27 Comparing aggregated tweet frequency during a NY weekend (2015-16 vs COVID)

Sentiment analysis of tweets from before and during COVID show interesting trends in customer behavior. While sentiment increases and decreases in level at similar rates before and during COVID, there is a slightly higher level of sentiment among public transit customers during COVID (Figure 28). This could have to do with low ridership numbers contributing to generally better performance by transit agencies across multiple metrics, leading to less complaint from customers. Additionally, customer sentiment was higher throughout the entire year during COVID as opposed to before COVID (Figure 29). This could be a result of, not necessarily transit-related tweets being made by active customer, but rather tweets being made in response to the transit authorities' efforts being made during the COVID pandemic. It is important to note however that during the summer months, tweet sentiment from the COVID pandemic did seem to normalize with pre-pandemic trends, as the first wave of COVID passed, and ridership climbed back up. By the holiday season, sentiment was up during COVID, and down pre-pandemic. It is possible that customers may have also been using happier language in their social media postings as a result of the strain that the pandemic had on customers throughout the year, with uncertainty of when it would conclude.

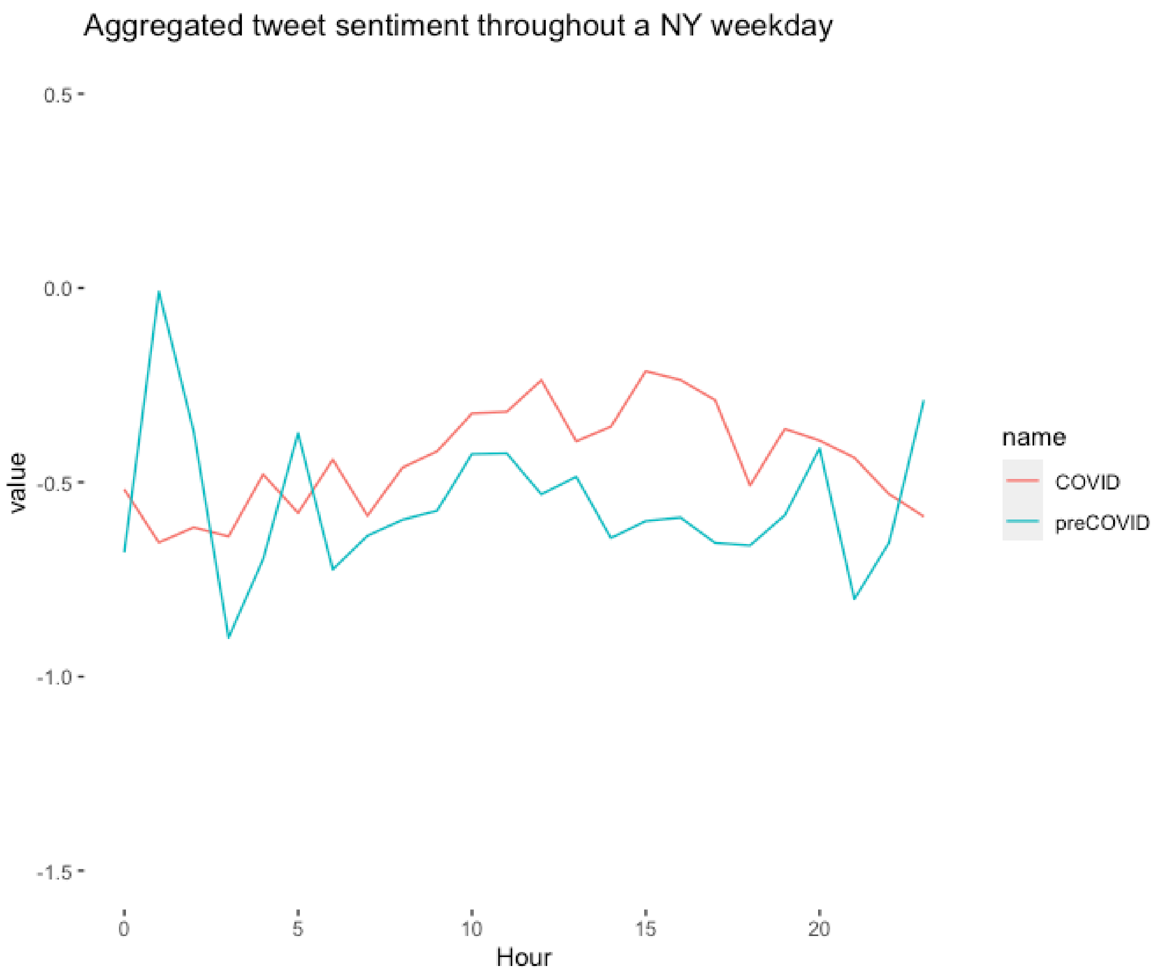

Figure 28 Comparing aggregated weekday tweet sentiment by hour (2015-16 vs COVID)

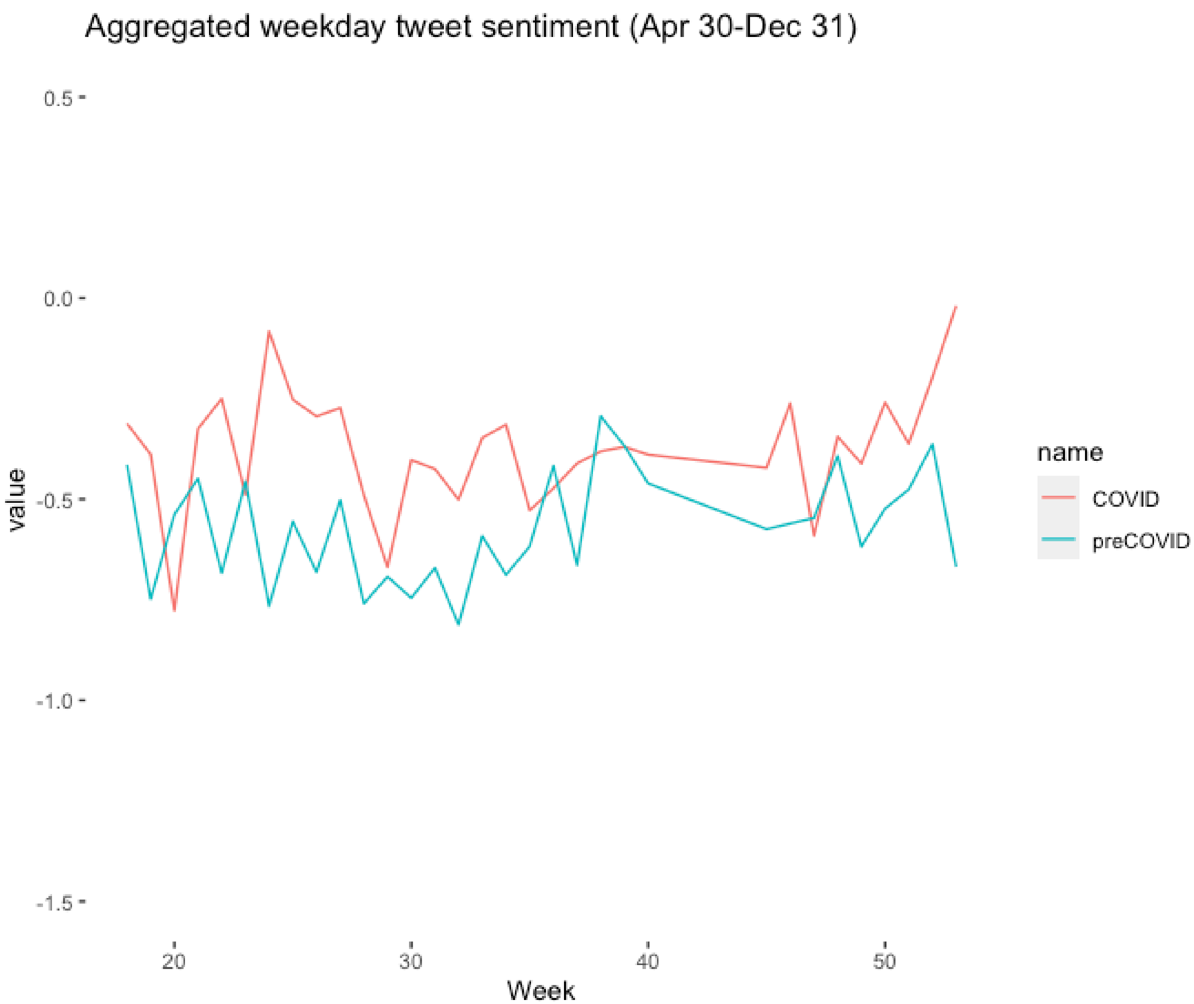

Figure 29 Comparing aggregated weekday tweet sentiment throughout the year (2015 v COVID)

Weekday sentiment levels were much lower than that of weekend sentiment levels; likely due to the fact that the work-week has concluded and social activities are the priority of many during this time. However, sentiment levels during the pandemic were consistently lower that pre-pandemic (Figures 30 and 31). Similarly, to during the weekday, weekend sentiment levels rose and fell at the same times throughout the day both before and during the pandemic; however, pandemic sentiment was much lower at each hour. The same trends apply to throughout the year where sentiment during the pandemic was consistently below that of pre-pandemic sentiment, even during the summer and holiday months. This may have to do with resentment and frustration towards the fact that at the beginning of the pandemic, social gatherings were being federally limited, businesses were limiting hours or closing entirely, and overall, it was not a safe environment for social interaction during the weekends. Additionally, MTA improvement projects were at an all-time high during off-peak hours and weekends because of decreased ridership, potentially contributing to further negative sentiment from customers. It is important to note that for all of the time periods analyzed, sentiment generally grew worse during major waves/spikes in COVID.

It is important to note that for all of the time periods analyzed, sentiment generally grew worse during major waves/spikes in COVID.

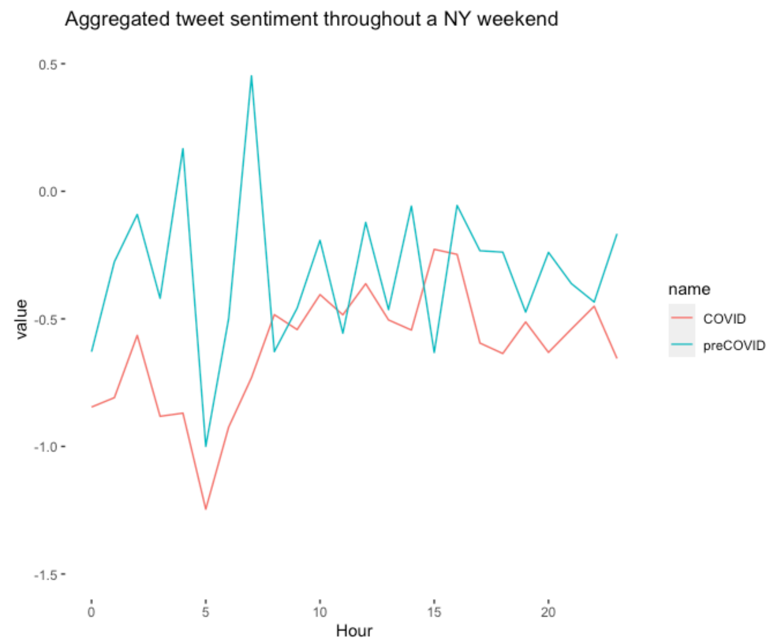

Figure 30 Comparing aggregated weekend tweet sentiment by hour (2015-16 vs COVID)

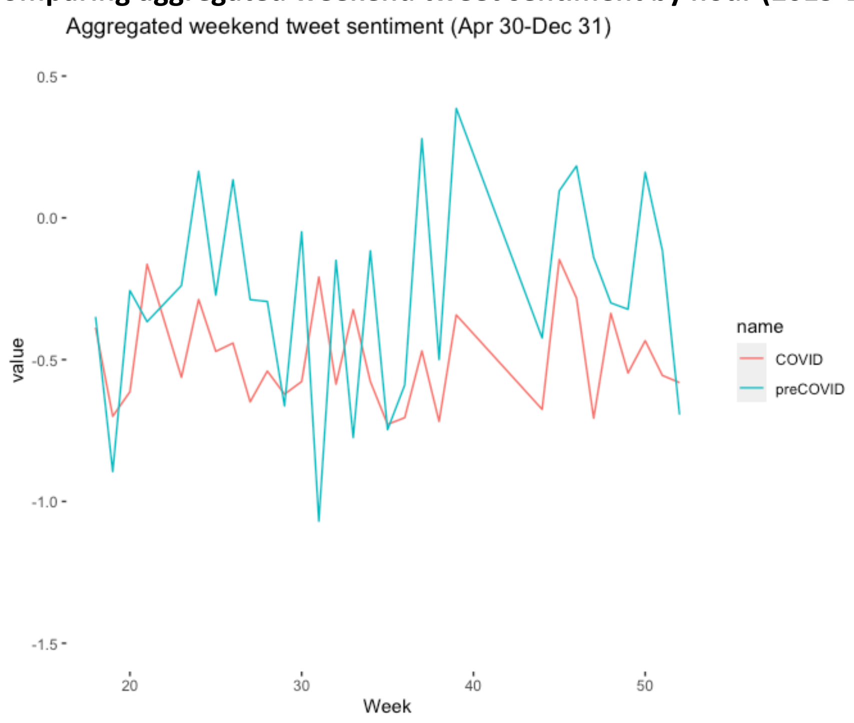

Figure 31 Comparing aggregated weekend tweet sentiment throughout the year (2015 v COVID)

Section 5 Event Data Analysis

The event data based on MTA's alerts have been analyzed for both pre-COVID and during COVID time period. Over months, construction is the major event category. Additionally, there are more there are many more special event alerts during COVID. The reduced service during COVID helped MTA address some maintenance issues. This is evidenced by a higher amount of track work events, particularly during June 2020.

Figure 32 Top-10 event categories over each month (2015 v COVID)

Figure 33 shows event type by time of day. Compared pre-COVID to during COVID, there are more operational activity and police activity events. This increase in pandemic-related alerts can also be seen in the comparative plot in Figure 34.

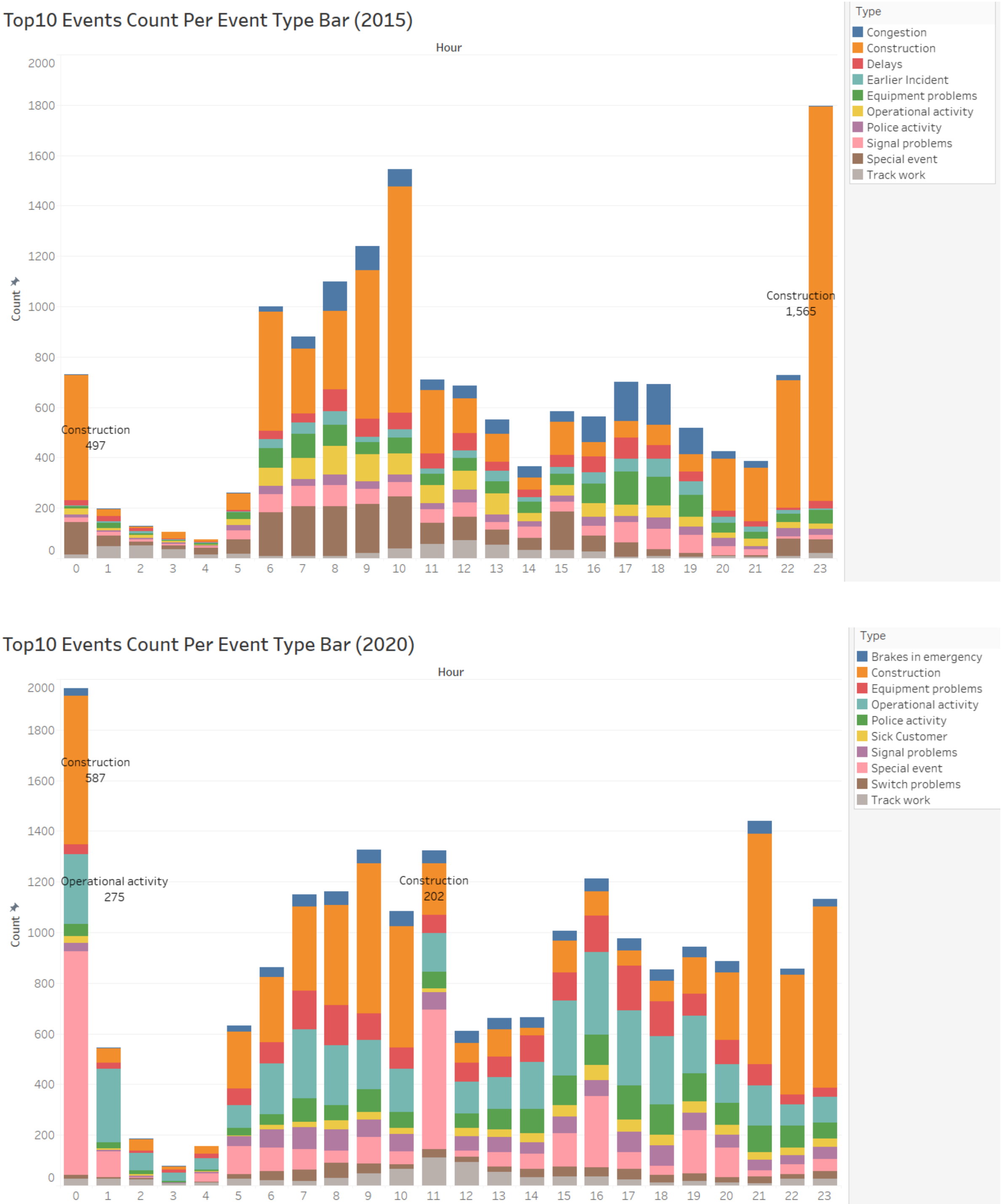

Figure 33 Top-10 event categories by time of day (2015 v COVID)

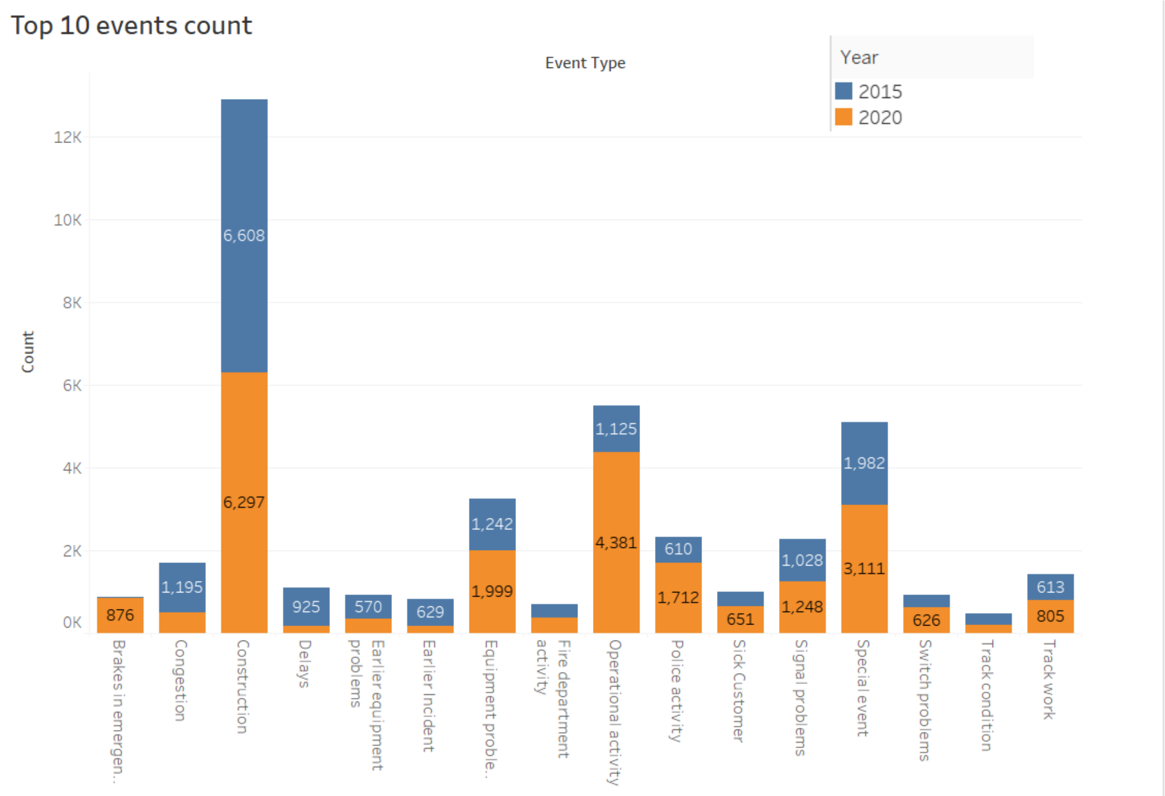

Figure 34 Top-10 event categories compared (2015 v COVID)

The content of police activity-related events was also analyzed and compared before & during the pandemic as shown in Figure 35. It can be see that there are much more service change-related content during the pandemic with phrases like "service to … from …", "to… from… skipped", "to… from… bypassed", "will start… end" occurring prominently.

The increase in operational activity & police activity events are mainly due to alerts coming up due to pandemic-related issues. The relative increase in pandemic-related and service-related topics can also be observed in transit-related tweets as noted in the previous section.

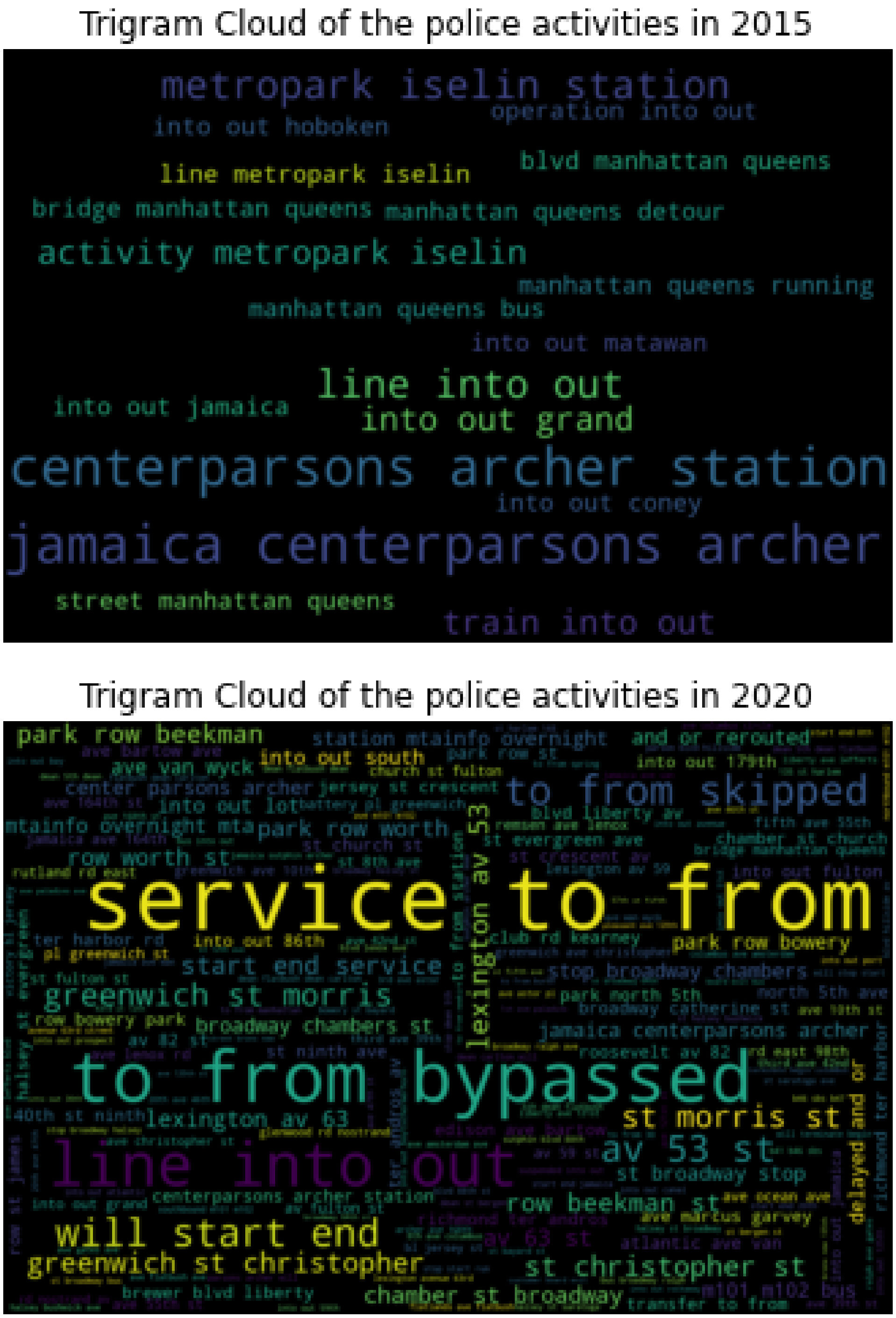

Figure 35 Comparison of text in police activity events (2015 v COVID)

There are higher number of constructions events during the AM peak period during COVID (as seen in Figure 36). This is largely due to the reduced demand and reduced service that the transit agencies were able to schedule more maintenance work during this time. A similar increase in operational activity including alerts related to the pandemic also during the peak period compared to pre-pandemic period.

Events Count Per Event Type (2015)

Hour

Type	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Congestion	2	1	2			1	21	47	116	94	68	42	51	57	45	40	101	158	162	105	31	27	20	4
Construction	497	28	8	26	7	69	471	258	311	591	898	252	137	111	48	134	58	66	79	70	205	213	506	###
Delays	23	23	12	2	2	4	35	36	86	72	67	59	70	38	27	47	63	84	54	37	23	21	10	30
Earlier Incident	1	5	9	1	4	2	34	44	55	21	32	21	28	40	18	27	44	50	72	54	26	21	14	6
Equipment problems	11	21	5	5	8	29	78	96	84	47	62	44	52	49	45	46	78	130	113	87	39	26	35	52
Operational activity	24	9	11	6	1	23	73	85	115	109	86	72	74	82	34	40	52	45	48	39	21	31	24	21
Police activity	12	7	7	7	5	21	34	25	42	31	29	25	50	31	20	24	36	26	45	33	32	11	31	26
Signal problems	17	13	10	6	6	36	69	82	83	58	57	54	59	30	46	39	40	80	83	73	36	25	10	16
Special event	130	43	15	14	26	59	174	196	196	195	206	82	93	61	49	153	63	56	26	15	3	4	68	55
Track work	14	48	50	37	15	17	10	10	10	21	39	58	71	53	32	33	27	7	9	5	9	7	10	21

Events Count Per Event Type (2020)

Hour

Type	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Brakes in emergency	30	3	1	1		24	38	50	54	54	60	51	47	44	42	40	50	48	45	41	46	52	25	30
Construction	587	57	47	14	29	223	256	333	397	591	478	202	78	108	29	125	98	59	78	145	264	910	473	716
Equipment problems	38	26	9	11	18	67	85	152	158	106	85	71	75	82	105	112	143	178	140	84	97	84	37	36
Operational activity	275	290	69	33	46	88	201	272	238	195	170	153	126	125	187	295	326	297	270	228	154	158	83	102
Police activity	49	23	16	6	7	31	42	92	60	88	63	66	57	79	93	117	119	135	118	111	85	103	89	63
Sick Customer	26	6	5		2	4	17	22	33	30	24	15	32	30	37	46	60	45	41	46	39	31	28	32
Signal problems	35	7	6	1	5	38	71	87	86	70	69	70	57	62	45	64	64	83	81	68	51	42	37	49
Special event	883	101	2		34	112	93	80	47	105	49	550	24	56	60	132	281	64	37	170	118	25	40	48
Switch problems	14	6	6		1	18	37	44	61	38	19	34	21	20	34	39	37	44	29	32	20	25	18	29
Track work	27	27	25	12	14	26	21	19	29	48	66	110	94	55	32	37	35	23	13	17	12	10	26	27

Figure 36 Event categories compared by time of day (2015 vs COVID)

Section 6 Key Findings

After analyzing the Twitter data, it is clear that customers are most likely to tweet about train delays and wait times of trains. Due to the nature of the New York City Subway system, with most trips often requiring at least one transfer, time is essential to the average New York commuter, where a delay in one train can further delay the arrival time of the customer. From the data collected between June 2015 and May 2016, the overall list of metrics customers expressed with concern on Twitter with the accuracy of the itinerary, timeliness of trains, station environment, traffic-related service, crowdedness of trains and stations, cleanliness, train speed, and lastly, fare cost. Negative sentiment outweighed positive sentiment by a huge margin, with only a small percentage of tweets pertaining to positive experience when riding transit. It was also discovered that tweets related to subway lines were nearly four to five times as much as those related to commuter rail lines. The No. 1, 2, and 7 trains were the most frequently tweeted, while the MNR Hudson and Harlem Lines, followed by the LIRR Port Washington Line were the top commuter rail lines mentioned on Twitter (Figure 11, 12).

Analyzing transit customer posts in Twitter during the pre-COVID and pandemic periods, it appears that tweets related to transit were more common prior to COVID as opposed to during the pandemic. During the pandemic, transit users were most concerned with the cleanliness of the system as they have expressed concerns regarding the health and safety issues of the transit system. After the start of the COVID-19 pandemic, there was a drastic change in the list and order of metrics which customers expressed with concern on Twitter. The cleanliness became the top concern just after the timeliness of trains. The crowdedness of trains also ranked higher than the station environment due to the uncertainty of COVID transmission throughout the first year of the pandemic. On weekdays, sentiment during the pandemic was generally higher than the pre-pandemic, but lower during the weekends. This change in sentiment may be a result from the change in service dynamics implemented by the MTA, including the reduction in ridership and the increase in improvement projects during off-peak hours and weekends. A similar pattern on the increase in service and operation-related events was also observed from the transit event-related data which the transit agencies publicly released. Additionally, changes in work culture may also play a role in how often customers tweet about service or their ability and willingness to tweet throughout the day when riding transit.

When compared with the performance metrics calculated by the MTA, it was clear that the difference in metrics mentioned in Twitter may be due to the discrepancy of on-time performance percentages between subway (< 70% accuracy) and commuter rail (90%+ accuracy) (Figure 13). There was no correlation between the performance of service and the term frequency of words throughout the year, as all the words maintained the same upsurges and decreases in term frequency during each month. Ranking of metrics also maintained the similar distribution throughout the year, with minimal deviation. Although the No. 1, 2 and 7 trains were the most mentioned train lines, they were also the first, third, and fourth highest performing lines respectively, based on the wait assessment percentages (Figure 11, 13). The mean distance between fleet failure, and injury rates among the separate MTA agencies were studied, but no correlation could be made to Twitter activity. Numbers for tweets regarding safety and mechanical failure were too scarce to make any concrete conclusions on the effects of performance within these metrics on customer feedback. When compared to different metrics, there was no clear explanation for the variation in customer tweet frequency reported during a particular month. Ridership specifically showed mixed correlation when compared to changes in tweet frequency. This suggests that customer behavior on Twitter may be affected by the time of year. Different social, weather, customer, and service patterns may also directly impact customer desire to share their transit experience on social media. Specifically, it is important to note that customer sentiment was at all-time lows, and tweet frequencies at all-time coincided with highs in MTA disrupted service reports during the same time intervals.

This study has proposed a methodology that utilize text and sentiment analysis of social media data to determine performance metrics customers frequently mentioned when riding the New York Transit system and how they compare to those reported by the agency. Mining data from social media could provide additional insights to customer perceptions about the quality of service to transit operators. The data from social media is readily available and the transit operators will not need to develop surveys to collect user's perceptions as obtaining statistically representative responses from surveys can be very challenging and usually requires tremendous resources.

Section 7 Future Research

This study explored the use of Twitter data as an alternative to surveys in conducting a transit analysis. While a variety of methods were implemented in this research, Twitter data contains a surplus of information that was not able to be fully investigated for this paper. Emoji and emoticon analysis is to be immediately researched next, as emojis act as the bridge between the user's emotions and their tweets. Emojis have a growing prevalence in Twitter and social media culture and are just as vital to sentiment analysis as words that carry sentiment. Separating data by its context can allow us to see which sector of transit users are most frequently mentioning, similarly to the monthly metric analysis. In addition to month-by-month analysis, it would be beneficial to explore week-by-week, day-by-day, and even hour-by-hour analysis to see how customer sentiment fluctuates throughout the week and day. It is common for MTA to conduct reduced or altered service on weekends due to construction and maintenance, which can yield radically different results than that of the average weekday. MTA service announcement data is available on a daily basis and could be used for this portion of research. It would also be essential to study how Twitter activity and customer expectations differ in the wake of New York-specific natural disasters, pandemics, and other unpredictable events. The most recent example where this study could be useful is with a situation like the COVID-19 pandemic, with customers and transit agencies holding a generally higher standard for cleanliness over frequent service than usual. NLP analysis of general, non-transit related tweets from pre-pandemic and during the pandemic could be useful for understanding the change in linguistics from these two time periods, and how that may play a role in how linguistics are used by public transit customers, and how that could further affect customer sentiment levels. Location analysis would be useful in helping to understand which neighborhoods, towns, or cities put out the most tweets and require the most attention. Inclusion of the B Division subway lines when determining which lines most customers have run into issues with would more greatly represent the customer population for New York City Transit, as these lines account for nearly two-thirds of New York City Subway service. Converting the code used in this research into an easy-to-use computer application could allow future researchers at The City College of New York to seamlessly continue this branch of research while simultaneously adapting it to the continuously evolving technology and research for years to come. Overall, expanding the representation of customers is the key objective moving forward with tidy text sentiment analysis. This will ensure a more accurate and fair assessment of customer sentiment across the New York metropolitan area, and transit systems worldwide.

Section 8 References

1. Ali, Farman, et al. "Fuzzy Ontology-Based Sentiment Analysis of Transportation and City Feature Reviews for Safe Traveling." Transportation Research Part C: Emerging Technologies, vol. 77, 2017, pp. 33–48., doi:10.1016/j.trc.2017.01.014.

2. Chakraborty, Pranamesh, and Anuj Sharma. "Public Opinion Analysis of the Transportation Policy Using Social Media Data: A Case Study on the Delhi Odd–Even Policy." Transportation in Developing Economies, vol. 5, no. 1, 2019, doi:10.1007/s40890-019-0074-8.

3. Collins, Craig, et al. "A Novel Transit Rider Satisfaction Metric: Rider Sentiments Measured from Online Social Media Data." Journal of Public Transportation, vol. 16, no. 2, 2013, pp. 21–45., doi:10.5038/2375-0901.16.2.2.

4. Delbosc, Alexa, and Graham Currie. "Modelling the Causes and Impacts of Personal Safety Perceptions on Public Transport Ridership." Transport Policy, vol. 24, 2012, pp. 302–309., doi:10.1016/j.tranpol.2012.09.009.

5. De Vos, J. (2018). Towards happy and healthy travellers: A research agenda. *Journal of Transport & Health*, *11*, 80-85.

6. El-Diraby, Tamer, et al. "Linking Social, Semantic and Sentiment Analyses to Support Modeling Transit Customers' Satisfaction: Towards Formal Study of Opinion Dynamics." Sustainable Cities and Society, vol. 49, 2019, p. 101578., doi:10.1016/j.scs.2019.101578.

7. Halvorsen, Anne, et al. "Passenger-Centric Performance Metrics for the New York City Subway." *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2673, no. 1, 2019, pp. 417–426., doi:10.1177/0361198118822297.

8. Hasan, Samiul, and Satish V. Ukkusuri. "Urban Activity Pattern Classification Using Topic Models from Online Geo-Location Data." Transportation Research Part C: Emerging Technologies, vol. 44, 2014, pp. 363–381., doi:10.1016/j.trc.2014.04.003.

9. Hoang, Thong, et al. "Crowdsensing and Analyzing Micro-Event Tweets for Public Transportation Insights." 2016 IEEE International Conference on Big Data (Big Data), 2016, doi:10.1109/bigdata.2016.7840845.

10. Houston, Douglas, and Thuy Luong. "Public Opinions of Light Rail Service in Los Angeles, an Analysis Using Twitter Data." 2015.

11. Ingvardson, Jesper Bláfoss, and Otto Anker Nielsen. "The Relationship between Norms,

Satisfaction and Public Transport Use: A Comparison across Six European Cities Using Structural Equation Modelling." Transportation Research Part A: Policy and Practice, vol. 126, 2019, pp. 37–57., doi:10.1016/j.tra.2019.05.016.

12. Levine, Brian, et al. "Measuring Subway Service Performance at New York City Transit: A Case Study Using Automated Train Supervision (ATS) Track- Occupancy Data." MTA New York City Transit, 2013.

13. Monmousseau, P., Marzuoli, A., Feron, E., & Delahaye, D. (2019). Passengers on social media: A real-time estimator of the state of the US air transportation system.

14. Nathanail, Eftihia. "Measuring the Quality of Service for Passengers on the Hellenic Railways." Transportation Research Part A: Policy and Practice, vol. 42, no. 1, 2008, pp. 48–66., doi:10.1016/j.tra.2007.06.006.

15. New York MTA Unveils Customer-Focused Subway Perfor mance Dashboard. Metro Magazine. http://www.metro- magazine.com/management-operations/news/725648/new- york-mta-unveils-customer-focused-subway-performance- dashboard. Accessed June 27, 2018.

16. Kamil Raczycki, Marcin Szymański,, Yahor Yeliseyenka and Piotr Szymański, Tomasz Kajdanowicz. 2021. Spatial Data Mining of Public Transport Incidents reported in Social Media. In Proceedings of SIGSPATIAL '21. ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145

17. Rashidi, Taha H., et al. "Exploring the Capacity of Social Media Data for Modelling Travel Behaviour: Opportunities and Challenges." Transportation Research Part C: Emerging Technologies, vol. 75, 2017, pp. 197–211., doi:10.1016/j.trc.2016.12.008.

18. Robinson, David, and Silge, Julia. Text Mining with R: A Tidy Approach, 2017.

19. "Social Media Use during Disasters." National Consortium for the Study of Terrorism and Responses to Terrorism, University of Maryland, 2012, www.start.umd.edu/sites/default/files/files/publications/START_SocialMediaUseduring Disasters_LitReview.pdf.

20. Tarte, L., B. Levine, A. Caspari, J. Fong, T. Huynh, and A. Reddy. Improving Transparency at the New York Metropolitan Transportation Authority Using Service Performance Dashboards. Presented at 98th Annual Meeting of the Transportation Research Board, Washington, D.C., 2019.

21. Transport for London. London Underground Performance Report. Transport for London, 2017.

22. Wińska, Monika et al. "Determining the level of satisfaction among users of public transport in Lublin." The Archives of Automotive Engineering – Archiwum Motoryzacji, vol. 85, no.3, 2019, pp. 19-39. doi:10.14669/AM.VOL85.ART2.

23. Zhang, Zhenhua, et al. “A Deep Learning Approach for Detecting Traffic Accidents from Social Media Data.” Transportation Research Part C: Emerging Technologies, vol. 86, 2018, pp.580–596, doi:10.1016/j.trc.2017.11.027.