

Twilytics: A Social Perception Analysis of Public Transit Systems during the COVID-19 Pandemic

Egbe-Etu Etu, Ph.D.
Imokhai Tenebe, Ph.D.
Ankur Parma

Likhitha Yelamanchili
Dang Minh Nhu Nguyen
Louis Tran

Ihor Markevych



Mineta Transportation Institute

Founded in 1991, the Mineta Transportation Institute (MTI), an organized research and training unit in partnership with the Lucas College and Graduate School of Business at San José State University (SJSU), increases mobility for all by improving the safety, efficiency, accessibility, and convenience of our nation's transportation system. Through research, education, workforce development, and technology transfer, we help create a connected world. MTI leads the [Mineta Consortium for Transportation Mobility \(MCTM\)](#) funded by the U.S. Department of Transportation and the [California State University Transportation Consortium \(CSUTC\)](#) funded by the State of California through Senate Bill 1. MTI focuses on three primary responsibilities:

Research

MTI conducts multi-disciplinary research focused on surface transportation that contributes to effective decision making. Research areas include: active transportation; planning and policy; security and counterterrorism; sustainable transportation and land use; transit and passenger rail; transportation engineering; transportation finance; transportation technology; and workforce and labor. MTI research publications undergo expert peer review to ensure the quality of the research.

Education and Workforce

To ensure the efficient movement of people and products, we must prepare a new cohort of transportation professionals who are ready to lead a more diverse, inclusive, and equitable transportation industry. To help achieve this, MTI sponsors a suite of workforce development and education opportunities. The Institute supports educational programs offered by the

Lucas Graduate School of Business: a Master of Science in Transportation Management, plus graduate certificates that include High-Speed and Intercity Rail Management and Transportation Security Management. These flexible programs offer live online classes so that working transportation professionals can pursue an advanced degree regardless of their location.

Information and Technology Transfer

MTI utilizes a diverse array of dissemination methods and media to ensure research results reach those responsible for managing change. These methods include publication, seminars, workshops, websites, social media, webinars, and other technology transfer mechanisms. Additionally, MTI promotes the availability of completed research to professional organizations and works to integrate the research findings into the graduate education program. MTI's extensive collection of transportation-related publications is integrated into San José State University's world-class Martin Luther King, Jr. Library.

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and accuracy of the information presented herein. This document is disseminated in the interest of information exchange. MTI's research is funded, partially or entirely, by grants from the California Department of Transportation, the California State University Office of the Chancellor, the U.S. Department of Homeland Security, and the U.S. Department of Transportation, who assume no liability for the contents or use thereof. This report does not constitute a standard specification, design standard, or regulation.

Twilytics: A Social Perception Analysis of Public Transit Systems during the COVID-19 Pandemic

Egbe-Etu Etu, Ph.D.

Imokhai Tenebe, Ph.D.

Ankur Parma

Likhitha Yelamanchili

Dang Minh Nhu Nguyen

Louis Tran

Ihor Markevych

September 2022

A publication of the
Mineta Transportation Institute
Created by Congress in 1991

College of Business
San José State University
San José, CA 95192-0219

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. 22-45	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle Twilytics: A Social Perception Analysis of Public Transit Systems during the COVID-19 Pandemic		5. Report Date September 2022	
		6. Performing Organization Code	
7. Authors Egbe-Etu Etu, PhD (ORCID #: 0000-0003-0457-6296) Imokhai Tenebe, PhD (ORCID #: 0000-0002-5862-6511) Ankur Parma (ORCID #: 0000-0003-3930-9542) Likhitha Yelamanchili (ORCID #: 0000-0002-5220-1184) Dang Minh Nhu Nguyen (ORCID #: 0000-0002-1025-1212) Louis Tran (ORCID #: 0000-0003-4237-0926) Ihor Markevych (ORCID #: 0000-0003-3891-6906)		8. Performing Organization Report CA-MTI-2210	
9. Performing Organization Name and Address Mineta Transportation Institute College of Business San José State University San José, CA 95192-0219		10. Work Unit No.	
		11. Contract or Grant No. 69A3551747127	
12. Sponsoring Agency Name and Address U.S. Department of Transportation Office of the Assistant Secretary for Research and Technology University Transportation Centers Program 1200 New Jersey Avenue, SE Washington, DC 20590		13. Type of Report and Period Covered	
		14. Sponsoring Agency Code	
15. Supplemental Notes			
16. Abstract In the United States, public transit ridership in 2020 declined by 79% compared to 2019 levels. With lockdowns implemented during the early days of the pandemic, direct human-to-human interactions migrated to virtual platforms (e.g., Facebook, Twitter, and Reddit). Social media platforms have aided researchers in answering numerous questions about current societal dilemmas, including COVID-19. This study investigates the public's perception of transit systems via a social media analysis given the emergence of vaccines and other COVID-19 preventive measures. Findings revealed themes of fear and confusion concerning the use of public transportation during the pandemic. The public had doubts regarding the vaccines' impact on transportation and movement throughout 2021, with most users concerned about the proliferation of new variants. Twitter users were concerned about the travel bans placed on African countries amidst the Omicron variant and urged the government to remove the bans. These findings will help bridge the gap between public health, transport, and commuter needs by helping transportation authorities and city planners better understand the social perception of transit systems during a pandemic.			
17. Key Words Public Transportation, Twitter, Text Mining, Social Media Mining, COVID-19		18. Distribution Statement No restrictions. This document is available to the public through The National Technical Information Service, Springfield, VA 22161.	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 36	22. Price

Copyright © 2022
by **Mineta Transportation Institute**

All rights reserved.

DOI: 10.31979/mti.2022.2210

Mineta Transportation Institute
College of Business
San José State University
San José, CA 95192-0219

Tel: (408) 924-7560
Fax: (408) 924-7565
Email: mineta-institute@sjsu.edu

transweb.sjsu.edu/research/2210

ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Transportation and Mineta Transportation Institute at San Jose State University (grant #: 69A3551747127).

CONTENTS

Acknowledgments	vi
List of Figures.....	ix
List of Tables.....	x
Executive Summary	1
1. Introduction.....	2
1.1 Study Motivation.....	2
1.2 Study objectives	3
1.3 Overview of the Report	3
2. Methods	4
2.1 Data Identification and Extraction.....	4
2.2 Data Preprocessing	5
2.3 Data Analysis.....	5
2.4 Outcomes	6
3. Results	8
3.1 Exploratory Results.....	8
3.2 Contextual Analysis Results.....	12
3.3 Sentiment and Spatial Analysis Results.....	14
3.4 Topic Modeling Results.....	16
4. Summary & Conclusions	20
Bibliography	23
About the Authors.....	25

LIST OF FIGURES

Figure 1. Proposed Transportation Tweet Analytics (Twilytics) Framework.....	4
Figure 2. Boxplot Showing the (A) Daily and (B) Monthly Number of Tweets in 2020	9
Figure 3. Boxplot Showing the (A) Daily and (B) Monthly Number of Tweets in 2021	10
Figure 4. An Analysis of Tweets for the Most Frequently Occurring Words. (A) Car and COVID-19; (B) Bus and COVID-19; (C) Public Transit and COVID-19; (D) Travel and Vaccine	12
Figure 5. Timeline of Tweet Frequency Related to (A) Public Transit, (B) Buses, (C) Cars, and (D) Air Transportation.....	13
Figure 6. Public Transit Sentiment Score and COVID-19 Infection Rate Across the US. (A) June to August 2020, (B) September to November 2020, (C) December 2020 to February 2021, (D) March to May 2021, (E) June to August 2021, And (F) September to November 2021	15
Figure 7. Public Transit Sentiment Score and Vaccination Rates Across the US. (A) December 2020 to March 2021, (B) April 2021 to June 2021, And (C) June 2021 to November 2021	16

LIST OF TABLES

Table 1. Daily Pairwise Comparisons Using Wilcoxon Rank-Sum Test	11
Table 2. Monthly Pairwise Comparisons Using Wilcoxon Rank-Sum Test	11
Table 3. Top Five Topics from Twitter for Each Keyword Obtained Through Topic Modeling	17

Executive Summary

In the United States, public transit ridership in 2020 declined by 79% compared to 2019 levels. With lockdowns implemented during the early days of the pandemic, direct human-to-human interactions migrated to virtual platforms (e.g., Facebook, Twitter, and Reddit). Social media platforms have aided researchers in answering numerous questions about current societal dilemmas, including COVID-19. This study investigates the public's perception of transit systems via a social media analysis given the emergence of vaccines and other COVID-19 preventive measures.

We developed a structured transportation tweet analytics framework (Twilytics) to analyze public discourse data (i.e., tweets from 2020 to 2021) on the impact of COVID-19 on transit systems. The framework has four main components. First, we extracted tweets between June 2020 to November 2021 from carefully curated keywords addressing transit services. Second, we pre-processed the data with data cleaning and feature engineering methods. Third, we performed descriptive and statistical analysis on the cleaned data. We hypothesized that the daily and monthly tweets related to transit systems will be significantly different. Lastly, we performed topic modeling to uncover the prominent themes of the public's perception of transit systems during the pandemic.

Overall, we extracted 44,320 tweets related to public transit in the US within the study period. Our results revealed that, on average, from June 2020 to November 2021, July (103) and Tuesdays (91) had the highest transit-related tweets. Kruskal-Wallis's analysis of variance test results showed a statistically significant difference ($p < 0.05$) in the number of transit-related tweets per month and day. The topic modeling findings revealed themes of fear and confusion concerning the use of public transportation during the pandemic. Second, the public had doubts regarding the vaccines' impact on transportation and movement throughout 2021, with most users concerned about the proliferation of new variants. Lastly, Twitter users were concerned about the travel bans placed on African countries amidst the Omicron variant and urged the government to remove the bans. These findings will help bridge the gap between public health, transport, and commuter needs by helping transportation authorities and city planners better understand the social perception of transit systems during a pandemic.

1. Introduction

Transportation is a key element of our society. The COVID-19 pandemic posed unique and unforeseen challenges to the movement of people, goods, and services globally. Out of fear of contracting the virus, most commuters stopped using public transportation in favor of driving their cars, riding bicycles, or walking (Habib & Anik, 2021). The disruptions caused by the pandemic have affected the movement of people in California, reducing the demand for public transportation. According to real-time ridership tracking data from TransitApp, an online technology platform that helps commuters navigate their cities, we observed a 33 percent decrease in demand as of November 28, 2021, in the San Francisco Bay Area (TransitApp, 2021). This decline poses a significant challenge for the US public transportation industry and more specifically for the Bay Area in California. In response, California, with an estimated population of 39.6 million, the government has been encouraging its residents to use public transportation by offering discounted or free transit for students, low-income workers, senior citizens, and employees through commuter programs such as Clipper Card to reduce congestion and environmental pollution. Despite this, the decreased usage of public transportation has led to social calamities as society functioned abnormally.

Facing calamities requires that people raise awareness and develop more positive solutions to improve current problems faced during these times. The COVID-19 pandemic has provided us with a practical case illustrating both the need to intervene in daily living to save lives and the need to restore normalcy to our lives in a healthy manner. Past research has shown that public transportation has been one of the hardest-hit industries during the pandemic (Astroza et al., 2020; Molloy et al., 2021). To investigate transportation-related issues during the pandemic, researchers have used methods such as survey questionnaires to collect data. As technology has advanced, direct human-to-human interaction as a method to harvest data has migrated to virtual platforms. This has made data collection from face-to-face surveys no longer a priority. Instead, people now express themselves through social media posts from various social media platforms such as Facebook, Instagram, Twitter, and Reddit. Approximately two-thirds of American adults, and nearly 3.8 billion people worldwide, use social networking sites (Arafat, 2020). This vast database has aided researchers in government and private institutions in answering numerous concerns about current societal dilemmas, including COVID-19. This study seeks to learn about the issues surrounding transportation through the lens of social media.

1.1 Study Motivation

In recent times, effective COVID-19 preventative methods (i.e., social distancing, wearing masks in public settings, and vaccines) have succeeded in limiting the transmission of the virus. However, there has been a slow return to the use of transit systems by commuters. For instance, reports from the US Department of Transportation (DOT) show a 30 to 50 percent decrease in the total number of passengers who ride intercity buses and Amtrak between January 2020 and November

2021 (BTS, 2022). Another example is the San Francisco Bay Area, which boast more than 2 million riders pre-pandemic but now has 800 thousand riders as of April 2022 (BART, 2020). Overall, the advent of vaccines has contributed to growth in the use of public transport but, given the slow growth rate and current threats to the environment and traffic safety, in-depth research on people's opinions about public transportation is required. This research will be helpful in understanding how to educate people on how they can use public transport more safely without the concern of contracting COVID-19.

Numerous studies have discussed the impact of the COVID-19 pandemic on transportation. Some provide an overview of public transportation and COVID-19, as well as many research issues and policy initiatives (Tirachini & Cats, 2020). In all the transportation studies related to COVID-19, authors believe that more research is needed to restore the safety and effectiveness of the public transit system during the pandemic (Gkiotsalitis & Cats, 2021). Despite these studies, there have been few studies in the literature that capture the following: public discourse during the early phase of the pandemic, such as stating whether or not public transit systems are risky during the pandemic; adverse effects of lockdown on people's health; insufficient space for cyclists on roads; and economic loss for transit companies (Osorio-Arjona, Horak, Svoboda, & García-Ruíz, 2021; Taleqani et al., 2021). We are now over 25 months into the pandemic, with multiple transit policies and strategies implemented. We need to ascertain their efficacy in order for transit agencies to be adaptive to changes caused by the virus. Therefore, there is an urgent need to investigate the public perception of transit systems, especially with new variants of COVID-19 such as Delta and Omicron, and increased infection and vaccination rates (i.e., as of January 27, 2022, approximately 63.6 percent of the total U.S. population is fully vaccinated (CDC, 2022)). The public's opinions and concerns will help us better understand the transport system's current challenges and opportunities. Social media analysis can be an effective tool to analyze the public's opinions on transportation.

1.2 Study objectives

The project aims to investigate how comfortable people are with using transit systems, given the emergence of vaccines and other COVID-19 prevention methods (e.g., use of face masks and hand sanitizers), via a social media analysis.

1.3 Overview of the Report

The remainder of the report is organized as follows: Chapter Two describes the methodology; Chapter Three presents the results; while Chapter Four summarizes the key findings, and suggests policy implementation and future research needs.

2. Methods

To address our research objectives, we proposed a structured transportation tweet analytics (Twilytics) framework to analyze public discourse data (i.e., Twitter data) on the impact of COVID-19 and various responses to it on the different modes of transportation (Figure 1). The proposed Twilytics framework is divided into four stages: data collection, data processing, data analysis, and outcomes.

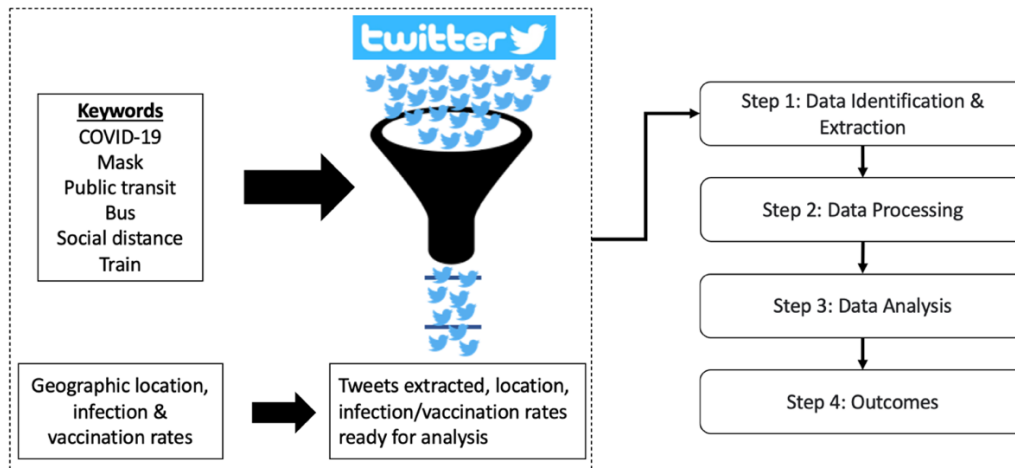


Figure 1. Proposed Transportation Tweet Analytics (Twilytics) Framework

2.1 Data Identification and Extraction

Tweets were collected from the Twitter API using a Python library called Tweepy (<https://docs.tweepy.org/en/stable/>). Tweepy is an easy-to-use python library for accessing Twitter’s Application Programming Interface (API). It requires the user to have a registered Twitter developer account. To extract the data, Tweepy requires an API key in the form of a consumer key, a consumer secret, an access token, and an access token secret, which are obtained from the Twitter developer account. Tweepy then queries the Twitter search API with user-defined keywords and stores the results of the search in JSON (JavaScript Object Notation) files or comma separated value (CSV) files. The extracted tweets include metadata information such as an identification number (ID), the date, time, and language. The tweets are also geotagged, which is an assigned geographic location of the Twitter user. The user’s location is extracted as latitude and longitude information. The geotagged tweet enables us to identify the Twitter user’s location, and extract infection and vaccination levels for that location to better understand the reason for their transportation-related tweet. In addition to the tweets, we extracted data on daily infection and vaccination rates across the United States, which are publicly available on the John Hopkins COVID-19 dashboard website.

Keywords used to extract the data were classified using the following categories: (a) public transit: “travel”, “train”, “bus”, “transit”, “rail”, “mobility”, “paratransit”; (b) COVID: “COVID”, “corona”,

“virus”, “pandemic”, “lockdown”, “social distance”, “community spread”; and (c) vaccination: “vaccine”, “Johnson”, “Pfizer”, and “Moderna”. These keywords allowed us to have a wide array of search terms and extract any tweet containing the terms. We extracted tweets between June 2020 and November 2021. This data collection covers the following periods: early pandemic; lockdowns; the pandemic’s peak; vaccine rollouts; and variants. Within this period, a total of 44,325 tweets were collected involving the above-mentioned keywords from all fifty US states.

2.2 Data Preprocessing

We processed and cleaned the tweets by removing punctuation signs, special characters, digits, emoticons, and hyperlinks so that the dataset contained only words. This was implemented using the Pandas library in Python. We removed retweets and any non-English tweets were excluded because translation was outside the scope of our study. In addition, we removed stop words (i.e., articles or prepositions) from the tweets.

2.3 Data Analysis

In this section, we discuss the methods required for analyzing the data beginning with data exploration, contextual analysis, and spatial analysis.

2.3.1 Exploratory and Statistical analysis

We performed exploratory analysis to visualize the daily and monthly tweets for our study period. Then, we performed a one-way analysis of variance (ANOVA) using the Kruskal-Wallis rank-sum test to investigate if there was any significant difference between the daily and monthly number of tweets related to public transit and COVID-19. We also conducted a multiple comparison test using the Wilcoxon rank-sum test with continuity correction to show which groups differed from each other with respect to tweet frequency. The Bonferroni method was used to adjust the p-values.

2.3.2 Contextual analysis

In this section, we analyzed the tweets using three approaches: text mining; sentiment analysis; and topic modeling.

- a) Text mining: Text mining is the process of analyzing large amounts of natural-language text to identify lexical and linguistic usage patterns of significance (Das, Sun, & Dutta, 2016). In text mining, a collection of large and structured text documents is represented by a corpus, which is then purified by removing redundant words, numbers, and punctuation. We used word clouds to visualize the results of the text mining analysis.
- b) Sentiment analysis: This uses natural language processing, text mining, and computational approaches to automate the classification of text into sentiments (Hussein, 2018; B. Liu,

2010; Zhang, Wang, & Liu, 2018). The method has been widely adopted in many fields including consumer information, marketing, websites, and social media. Sentiment analysis is an important tool for decision-making today. For this study, sentiment analysis was used to analyze the transportation tweets, and examine the scores of sentiments. This process gives us an opportunity to explore the mindset of Twitter users regarding COVID-19 and public transportation. We classified the tweets in the text corpus according to whether they expressed positive and negative feelings.

- c) Topic modeling: This is a method used for extracting the abstract topics (i.e., word or phrase patterns) that occur in a collection of documents (Liu, Tang, Dong, Yao, & Zhou, 2016). In this project, we adopted the Latent Dirichlet Allocation (LDA) approach (Blei, Ng, & Jordan, 2003). LDA is a method used for topic modeling to classify the text in unstructured documents. It builds a *topic per document* model and *words per topic* model, each modeled as Dirichlet distributions. The LDA algorithm identifies topics in unstructured data based on the word frequency from a set of documents, where topics are composed of a weighted list of words. The words are then categorized into topics, and these are named using frequently used words combined into a sentence.

The most prominent keywords, based on different events during the pandemic (i.e., early pandemic stage [March 11, 2020 to December 13, 2020], first vaccine dose [December 14, 2020 to January 15, 2021], second vaccine dose [January 16, 2021 to March 14, 2021], Delta variant [March 15, 2021 to November 22, 2021], booster shot [November 19, 2021], and Omicron variant [November 22, 2021]), help to understand the public's discourse on public transportation. Additionally, we analyzed the number of tweets related to different modes of transportation with respect to time. Using Python, we created an extra feature (Boolean value) for verifying the existence of keywords such as “coronavirus”, “train”, “transit”, “airport”, “car”, and “bus”.

2.3.3 Spatial analysis:

We mapped the cleaned results from the contextual analysis to a cleaned location and time data for the spatial analysis. Each extracted tweet is associated with a geographical location ID, geographic coordinates, city, state, and country. The analysis in this section is focused on exploring the sentiments based on (1) infection, and (2) vaccination rates from a cluster of tweet geolocations, since this affects the type of tweets.

2.4 Outcomes

The potential results of our proposed Twilytics framework include: (a) emerging themes based on timelines (which include challenges faced by commuters); (b) prominent themes by location, overlaid on maps, which will show commuters' perceptions, and infection/vaccination rates within the Bay Area; (c) sentiment analysis mapped on themes and topics to capture public perception of transit systems (i.e., current problems, solutions, and opportunities); (d) recommendations to aid

transit policymakers, planners, and transportation authorities in the development of guidelines that will boost and improve the use of public transit systems and restore the systems' ability to fulfill their societal role.

3. Results

3.1 Exploratory Results

The boxplot gives a quick visual assessment of the daily and monthly number of transportation-related tweets during the pandemic (2020 to 2021). For 2020, we analyzed transportation-related tweets for the last six months of the pandemic (June to December 2020). As illustrated in Figure 2a, we observed that tweets made on weekdays involved more transportation discussion when compared to tweets made during weekends. It was also observed that Wednesday had the highest number (128) of transportation-related tweets by Twitter users. Figure 2b also showed the monthly distribution of transportation-related tweets using a boxplot for the year 2020. July had the highest median of 146 tweets related to public transport, indicating that Twitter users were active and concerned during the early days of the COVID-19 outbreak.

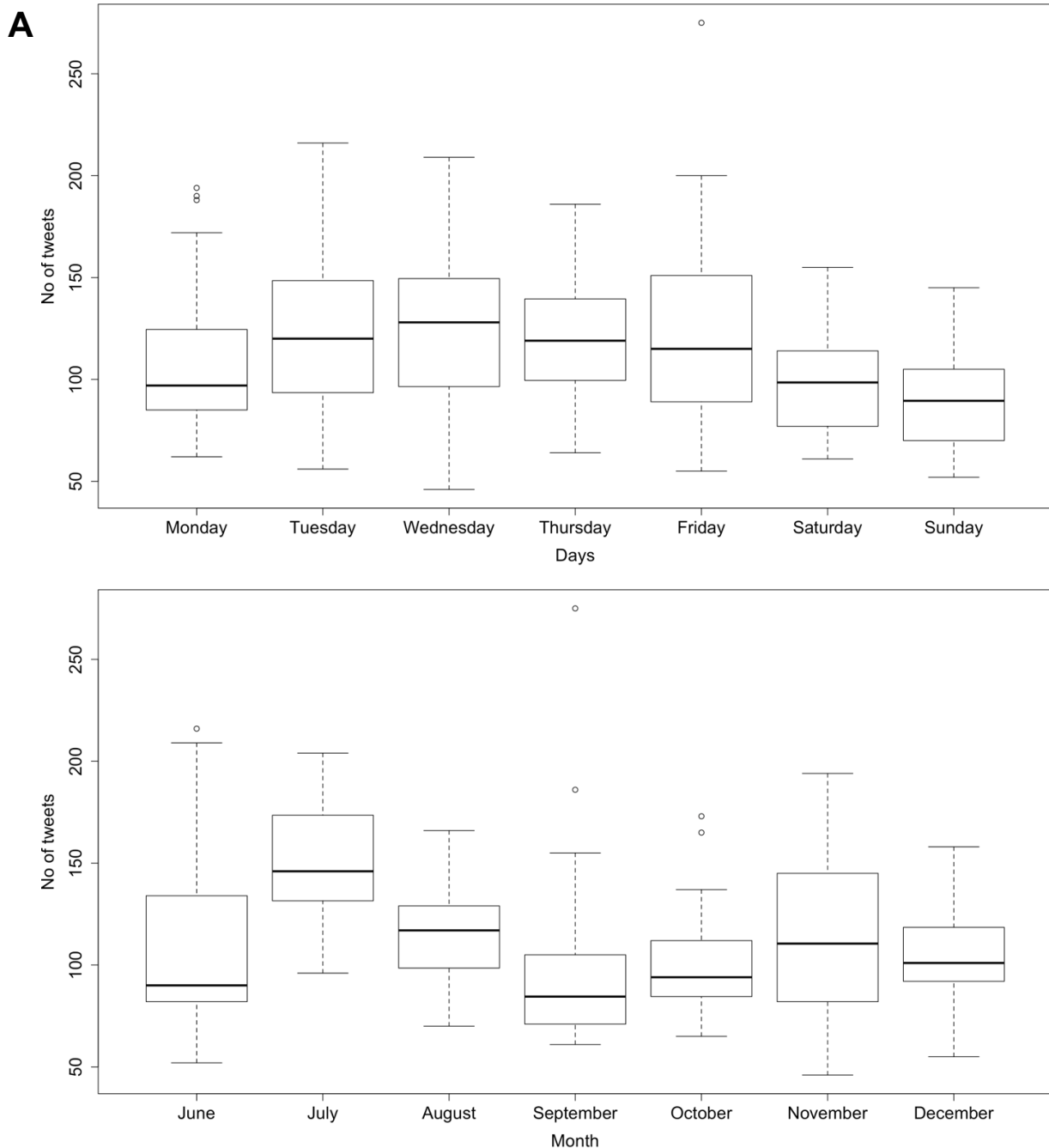


Figure 2. Boxplot Showing the (A) Daily and (B) Monthly Number of Tweets in 2020

Similar analysis was performed for the year 2021 and it was observed that most of the Twitter users were active during the start of the week (Mondays and Tuesdays) and in contrast users were less active during the weekends (refer to Figure 3a). The median value for daily transport related tweets was 63.5 (Tuesdays) and this was less than the previous year by a factor of 50.4. One reason for this reduction is that researchers, scientists, and public health officials who studied the virus in the early stages understood how it spread, and could give better insights to transit officials and the

general public for informed decision-making in 2021. Second, the reduction in tweets can be attributed to a gradual return to using public transit systems after the implementation of multiple safety measures including, but not limited to, social distancing, the use of hand sanitizers, the wearing of face coverings, and vaccines.

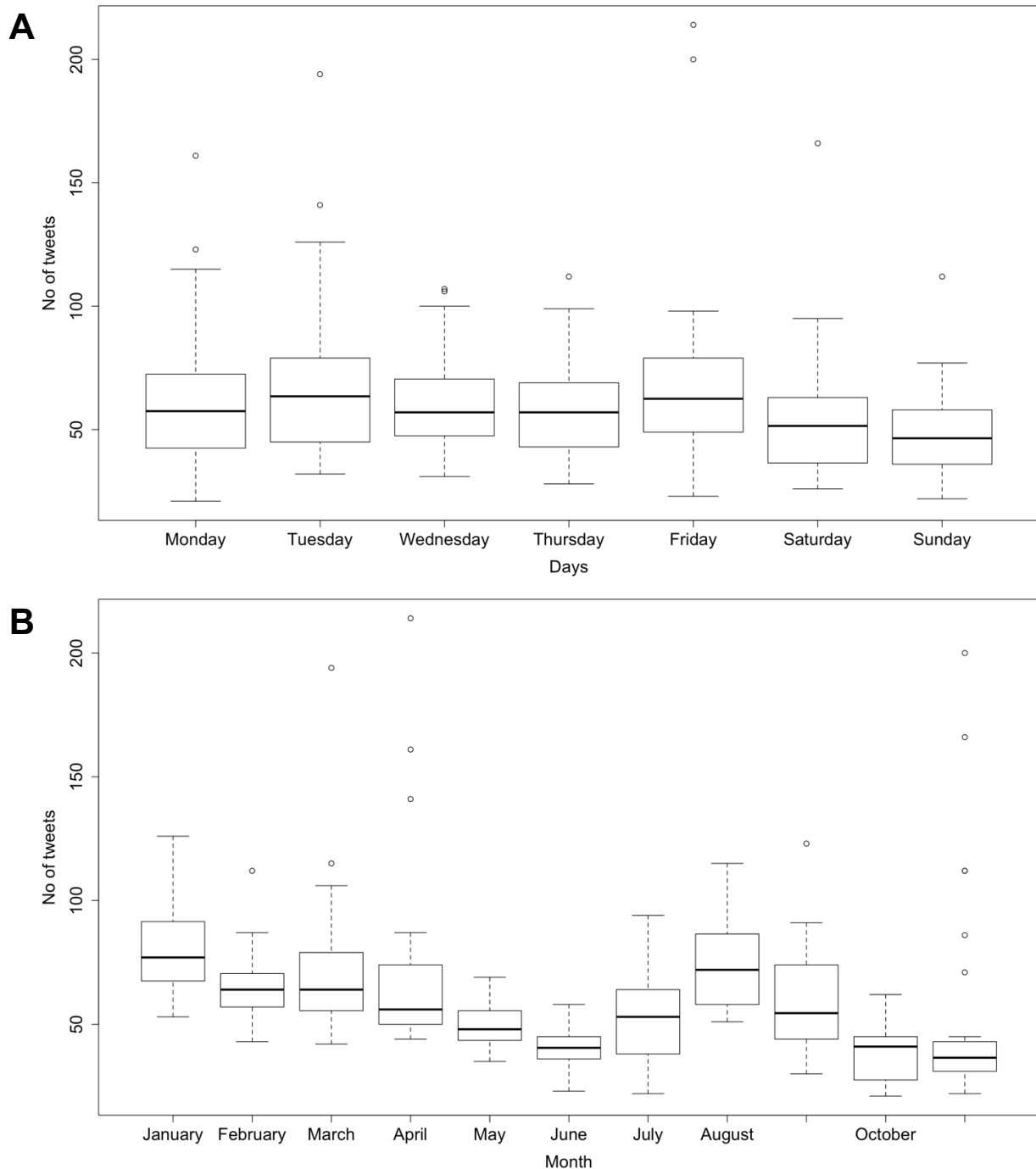


Figure 3. Boxplot Showing the (A) Daily and (B) Monthly Number of Tweets in 2021

The results in Figure 3b revealed that a greater number of tweets (median = 77) were generated during the months of January 2021 when the first COVID-19 vaccines were administered. In

addition, we found an increase in the number of tweets (median = 72) in August due to the Delta variant. This suggested that Twitter users were more actively tweeting about public transportation when new variants were found, and when vaccinations were rolled out.

The ANOVA test showed there was a significant difference between the number of tweets per day ($p < 0.05$). Table 1 presents the multiple comparison results, which shows a statistically significant difference between tweet frequency and days.

Table 1. Daily Pairwise Comparisons Using Wilcoxon Rank-Sum Test

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
Tuesday	0.457	-	-	-	-	-
Wednesday	0.790	0.781	-	-	-	-
Thursday	0.825	0.650	0.825	-	-	-
Friday	0.732	0.825	0.825	0.790	-	-
Saturday	0.229	0.039*	0.129	0.155	0.099	-
Sunday	0.024*	0.004*	0.012*	0.016*	0.007*	0.380

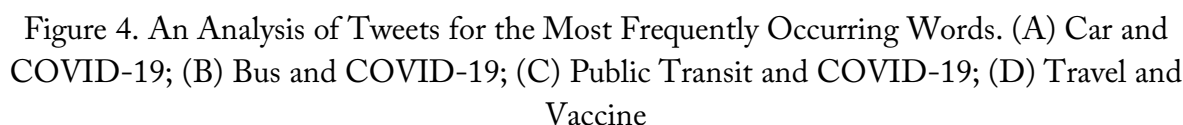
Note: *P-value* adjustment method: Bonferroni. Asterisk (*): Pairwise comparison is significant at the 0.05 level
The ANOVA test revealed a significant difference between the number of tweets per month ($p < 0.05$) and the multiple comparison test results showed a statistically significant difference between the number of tweets per month (Table 2).

Table 2. Monthly Pairwise Comparisons Using Wilcoxon Rank-Sum Test

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov
Feb	0.001*	-	-	-	-	-	-	-	-	-	-
Mar	0.043*	0.726	-	-	-	-	-	-	-	-	-
Apr	0.000*	0.160	0.137	-	-	-	-	-	-	-	-
May	0.000*	0.000*	0.000*	0.005*	-	-	-	-	-	-	-
Jun	0.052	0.485	0.227	0.621	0.238	-	-	-	-	-	-
Jul	0.382	0.071	0.134	0.052	0.000*	0.016*	-	-	-	-	-
Aug	0.043*	0.000*	0.000*	0.000*	0.000*	0.000*	0.750	-	-	-	-
Sept	0.238	0.166	0.634	0.106	0.000*	0.160	0.117	0.001*	-	-	-
Oct	0.130	0.750	0.399	0.747	0.160	0.947	0.003*	0.000*	0.240	-	-
Nov	0.656	0.750	0.947	0.905	0.160	0.750	0.053	0.071	0.749	0.444	-
Dec	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.747	0.130	0.000*	0.000*	0.027*

Note: *P-value* adjustment method: Bonferroni. Asterisk (*): Pairwise comparison is significant at the 0.05 level

After cleaning and processing the Twitter data, we used four corpora (i.e., based on a set of keywords) namely “car and COVID-19,” “bus and COVID-19,” “public transit and COVID-19,” and “travel and vaccine” to generate the word clouds as presented in Figure 4. Larger-sized words indicated more frequent occurrences on Twitter, which suggest that users discussed them more often.



MINETA TRANSPORTATION INSTITUTE

for wearing masks; and the need for kids to be vaccinated before using school buses. There was mixed messaging around the use of masks and their effectiveness. This frequency of messaging interfered with the ability of the government to reduce the spread of the virus and resulted in people looking for alternative modes of transportation.

Figure 4c represents the word cloud derived from the tweets containing the “public transit and COVID-19” keywords. The highly occurring words in these tweets were “transit,” “COVID,” “people,” “bus,” “delay,” “services,” “mask,” “workers,” “shortage,” and “vaccines.” To understand the broader context in the use of these words, tweets containing these words were analyzed. Twitter users expressed concern about the slow pace of trains during the pandemic, people being stranded due to reduced transportation options, and no trust/confidence in public transportation because of mixed messaging. Figure 4d shows that the most frequently used words in “travel and vaccine” related tweets were, “mandate,” “passport,” “COVID,” “mask,” “people,” “ban,” “vaccination,” “country,” and “air.” Tweets containing these words were extracted and evaluated. We observed the following among Twitter users: they encouraged people to take the vaccine for travel to resume as this was hurting the economy; concerns were raised on an air travel ban and that the ban created a bias towards African nations and the need for creating vaccine mandates for travel.

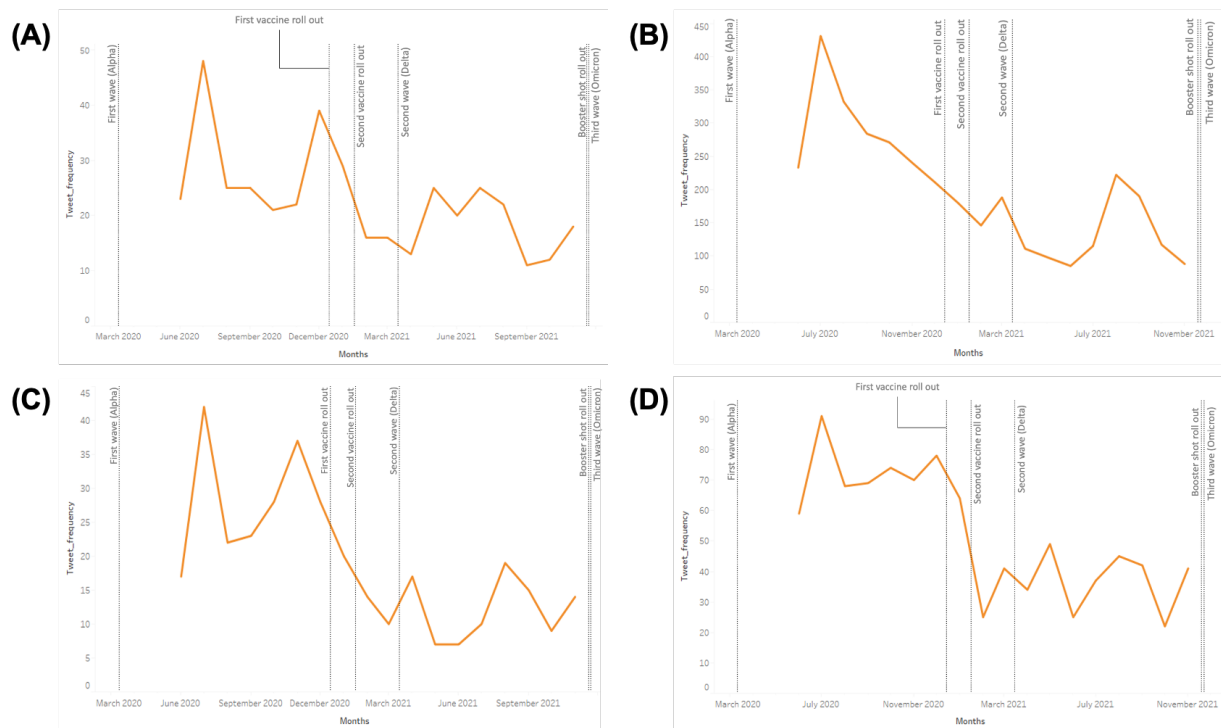


Figure 5. Timeline of Tweet Frequency Related to (A) Public Transit, (B) Buses, (C) Cars, and (D) Air Transportation

Therefore, based on the corpora, we focused on two main keywords and investigated their tweet frequency during the pandemic: public transportation; and modes of transportation (e.g., bus, car,

and air). The frequency of tweets indicates how interested users are in those keywords. We visualized the tweet frequency over different timelines during the pandemic (Figure 5). Given the importance of “public transportation” in our society and the effects of the pandemic on public transportation systems, we initially represented the frequency of tweets on the subject. Next, we looked into how frequently tweets about the three forms of transportation in the US—bus, car, and air—occur.

Tweets on public transportation more than doubled during the initial COVID-19 infection wave, as seen in Figure 5a. During these periods, we find shifts in tweet frequency. After the second vaccine rolled out, there was a decrease in the number of tweets referencing public transportation, which may suggest a rise in confidence to utilize public transportation if one is vaccinated. The emergence of the Delta variant resulted in an upsurge in tweets about public transportation. This implies that the public was efficiently using social media platforms to express their worries and concerns during the pandemic. Figure 5b showed a sharp increase in the number of tweets related to “buses” during the first COVID-19 wave between March 2020 and July 2020. A gradual decline in the bus related tweets can be observed from August 2020 as lockdowns and other social distancing measures were implemented. Furthermore, Figure 5b indicates a fluctuation about the topic on Twitter after the second vaccination was introduced in the US.

The highest number of tweets were observed during the first wave, and we see that the frequency of tweets increased along with the infection rates, despite the fluctuations after the second vaccine rollout (Figure 5c). Figure 5d depicts the frequency distribution of tweets about “air” transportation. The number of tweets fluctuated during the study period. The highest monthly tweet volume was in July 2020, and the lowest monthly tweet volume was in October 2021.

3.3 Sentiment and Spatial Analysis Results

As the second phase of the contextual analysis, sentiment analysis was performed on the cleaned Twitter data. The results of the sentiment analysis were merged with the spatial analysis to reveal the infection and vaccination rates of states across the US.

Part A: Infection rate vs sentiment score distribution

We merged the daily COVID-19 infection rate data from different states in the US with the sentiment scores of daily tweets to determine the public’s perception of the impact of COVID-19 on public transportation (Figure 6a-f). From June 2020 to August 2020, we visualized the number of confirmed infections and the sentiment scores for public transportation, as shown in Figure 6a. Notably, New York had a high infection rate, and 1,503 tweets were related to “public transportation and COVID or Coronavirus outbreak,” with a negative sentiment score of 36.06 percent. As the disease spread across the US, California had the highest confirmed COVID-19 infection rate during the second time frame (September to November 2020) (see Figure 6b). The number of public transportation tweets generated during this time period was 1,282 tweets, with

39.24 percent of the tweets having a negative sentiment score, indicating that Twitter users were concerned about public transit in the state of California and its contribution to the spread of the virus. Figure 6c and 6d show that California remained the state with the highest infection rate, though a slight decrease in the negative sentiment score was observed. The emergence of the Delta variant resulted in an increase in COVID-19 infections across the United States and we observed a modest rise in the number of infections from June 2021 to November 2021. A minor portion of tweets on public transportation have a negative sentiment score, as seen in Figure 6e and 6f, but California continued to have the highest number of infections.

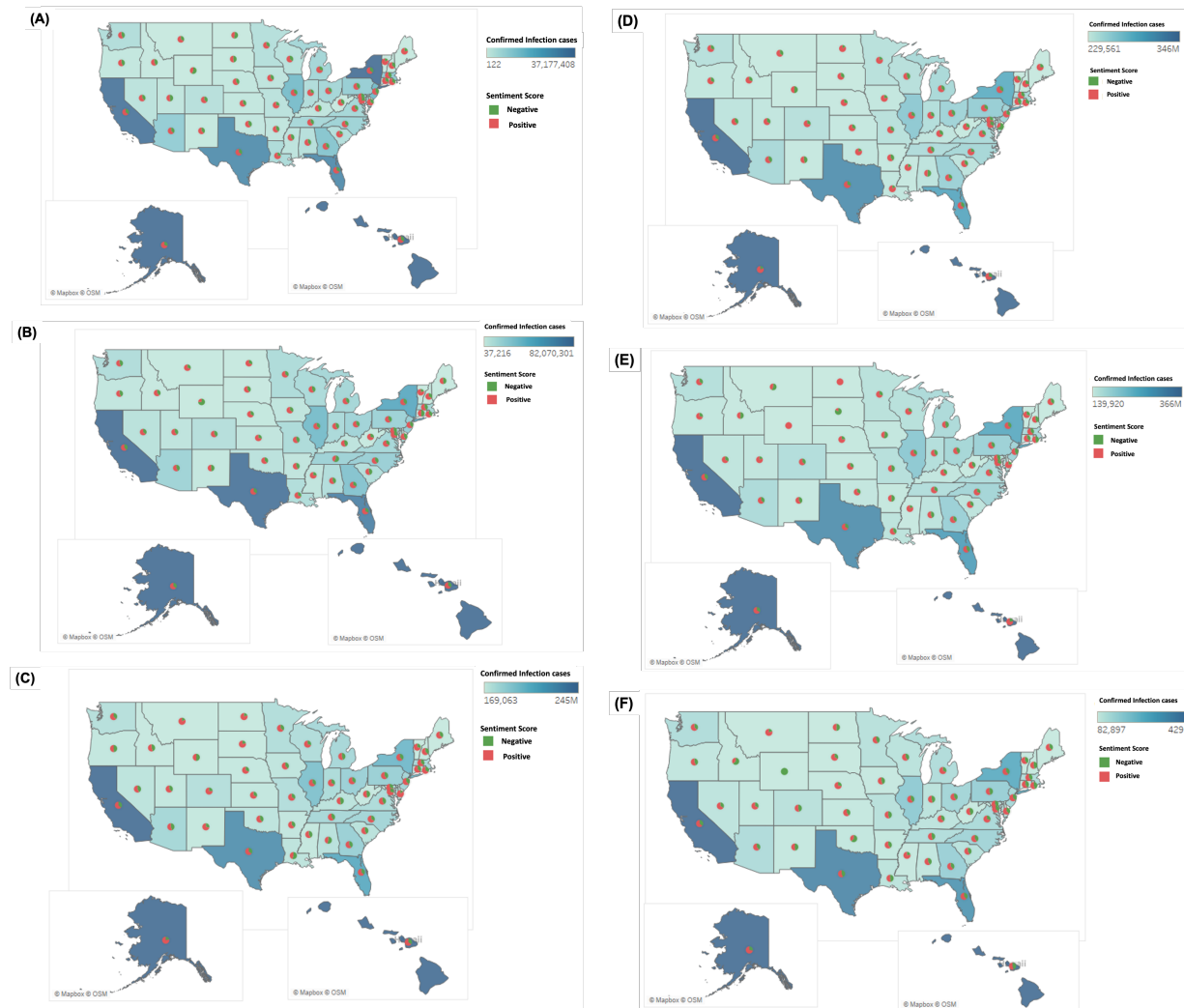


Figure 6. Public Transit Sentiment Score and COVID-19 Infection Rate Across the US. (A) June to August 2020, (B) September to November 2020, (C) December 2020 to February 2021, (D) March to May 2021, (E) June to August 2021, And (F) September to November 2021

Part B: Vaccination rate vs sentiment score distribution

Vaccinations were made available to the public in the United States starting in December 2020. The vaccination data from each state in the United States was retrieved and used in our study to examine the COVID-19 vaccination rates as well as the public transportation sentiment score distribution. The vaccination data and tweets were evaluated on a three-month basis (Figure 7).

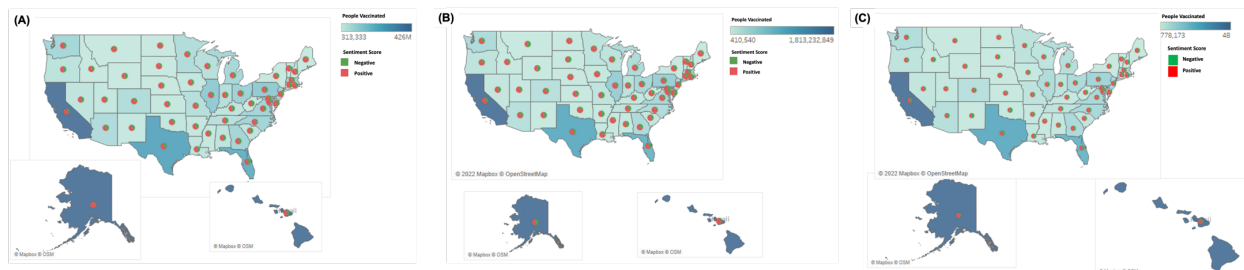


Figure 7. Public Transit Sentiment Score and Vaccination Rates Across the US. (A) December 2020 to March 2021, (B) April 2021 to June 2021, And (C) June 2021 to November 2021

From Figure 7a-c, it was observed that California had the highest vaccination rate, followed by Texas. It was also observed that most of the tweets generated during this period in all the quarters had a positive sentiment score. This implied that Twitter users' confidence in public transportation increased as the COVID-19 vaccines were accepted by citizens and administered across the country. Furthermore, results in Figure 7a-c are correlated with Figure 5a-d, which reveals a decrease in tweet frequency related to public transportation after vaccines were administered to citizens in record numbers.

3.4 Topic Modeling Results

As the final phase of the contextual analysis, topic modeling was performed on the processed Twitter data. The most prominent keywords with respective probabilities associated with different topics were extracted and listed in a chronological order based on the events that occurred during the pandemic (Table 3). In order to gain a full understanding of these co-occurring words, we reviewed sentences containing these word clusters from the processed tweets. Topics without any meaning were not listed or included in the result.

Table 3. Top Five Topics from Twitter for Each Keyword Obtained Through Topic Modeling

Topic number	Topic word clusters and their associated probabilities
Early pandemic	
1	Bus (0.011), COVID (0.011), Pandemic (0.010), Trump (0.009), People (0.006), Time (0.006), Train (0.005), Would (0.005), See (0.004), Quarantine (0.004)
2	Bus (0.017), Mask (0.015), COVID (0.014), Train (0.011), Wear (0.009), Masks (0.008), One (0.008), Wearing (0.007), People (0.006), Home (0.006)
3	COVID (0.022), People (0.011), Get (0.010), Pandemic (0.009), Virus (0.008), Like (0.008), Train (0.006), Vaccine (0.006), Transit (0.005)
4	COVID (0.021), Train (0.008), People (0.007), Pandemic (0.007), Vaccine (0.005), Trump (0.005), Country (0.005), China (0.004), Bus (0.004)
5	COVID (0.031), Pandemic (0.008), New (0.008), Quarantine (0.008), Days (0.007), Would (0.006), Thanksgiving (0.005), Holiday (0.005), Test (0.004), home (0.004)
First vaccine dose	
1	Get (0.009), Pandemic (0.008), People (0.007), Vaccine (0.004), Time (0.004), One (0.004), Work (0.004), Mask (0.003), Transit (0.003)
2	Train (0.009), Pandemic (0.008), Get (0.005), Vaccine (0.005), Go (0.004), Need (0.003), Still (0.003), Even (0.003), Bus (0.003)
3	Bus (0.009), Pandemic (0.007), Vaccine (0.007), New (0.007), People (0.006), Get (0.005), Train (0.005), Like (0.005), Home (0.005), Mask (0.004)
4	Pandemic (0.008), Bus (0.008), Train (0.008), Mask (0.004), Like (0.004), Test (0.004), Vaccine (0.004), Want (0.004), People (0.004)
5	People (0.011), Pandemic (0.010), Get (0.008), Vaccine (0.006), Train (0.006), Like (0.005), Year (0.005), Masks (0.004), Know (0.004)
Second vaccine dose	
1	Bus (0.010), Train (0.008), Home (0.007), Vaccine (0.006), People (0.005), Please (0.005), Want (0.005), Due (0.004), Test (0.004), Time (0.004)
2	Would (0.008), Bus (0.007), Trump (0.005), Could (0.005), Mask (0.005), Train (0.005), Take (0.004), Vaccine (0.004), Wear (0.004)
3	Vaccine (0.011), People (0.008), Bus (0.007), train (0.005), Vaccines (0.005), Day (0.005), Like (0.005), Would (0.005), Think (0.005), Year (0.004)
4	People (0.010), Train (0.007), Vaccine (0.007), Bus (0.007), Vaccinated (0.005), State (0.004), Way (0.004), Back (0.004), Round (0.004), Right (0.003)
5	Mask (0.009), Bus (0.008), Vaccine (0.008), Transit (0.006), People (0.005), Got (0.004), Train (0.004), Time (0.004), Today (0.004), Work (0.004)
Delta variant	
1	Mask (0.017), Bus (0.016), Train (0.009), People (0.009), Wear (0.008), Wearing (0.008),

Topic number	Topic word clusters and their associated probabilities
	Masks (0.007), Still (0.006), Pandemic (0.006)
2	Pandemic (0.010), Vaccinated (0.008), Get (0.008), Bus (0.008), People (0.007), Last (0.005), Time (0.005), Year (0.005), Train (0.004), Transit (0.004)
3	Train (0.016), Get (0.012), Bus (0.009), Vaccinated (0.008), Vaccine (0.007), Pandemic (0.007), Mask (0.006), Work (0.006), People (0.005), Time (0.005)
4	Vaccinated (0.011), Vaccine (0.009), School (0.007), Get (0.006), Vaccines (0.005), Bus (0.005), People (0.005), Fully (0.005), International (0.004)
5	Vaccine (0.009), Test (0.007), Bus (0.007), Vaccinated (0.006), Get (0.006), People (0.006), Need (0.006), Vaccination (0.005)
Booster shot and Omicron variant	
1	New (0.007), Variant (0.007), Train (0.005), Vaccinated (0.005), Back (0.005), Restrictions (0.005), Take (0.004), Ban (0.004), Africa (0.004), Test (0.004)
2	New (0.012), Variant (0.010), Countries (0.009), Bans (0.008), Africa (0.008), Vaccine (0.006), Restrictions (0.006), South (0.005), Biden (0.004)
3	Omicron (0.010), Variant (0.009), New (0.008), Africa (0.007), People (0.007), Vaccinated (0.006), Bans (0.005), Still (0.005), Ban (0.005), South (0.005)
4	Ban (0.010), Variant (0.009), Trump (0.008), Pandemic (0.007), Countries (0.007), People (0.007), Africa (0.007), New (0.006), Biden (0.006), Time (0.005)
5	Mask (0.006), Omicron (0.006), Variant (0.005), Bus (0.005), Train (0.005), Got (0.005), Vaccine (0.004), Pandemic (0.004), People (0.004), New (0.004)

Topic modeling on public transportation tweets during the early pandemic phase was conducted, and the results are shown in Table 3. It revealed that from the early pandemic results, the words with the highest probabilities in Topic 1 are “bus” (0.011), “COVID” (0.011), and “pandemic” (0.010), indicating that the public worried on how they were going to avoid the virus while using bus transportation. The words with small probabilities were “train” (0.005), “would” (0.005), “see” (0.004), and “quarantine” (0.004), expressing that the public were in support of infected people needing to quarantine and avoiding the use of public transit. In addition, Topic 2 revealed that the public supported the idea of commuters wearing a “face mask” (0.015) while taking a “bus” (0.017) or “train” (0.011). The word clusters in the third topic focus on people getting the vaccines to avoid contracting the virus while using public transportation. In Topic 4, Twitter users were focused on public transportation within the “US” (0.007) and other countries like “China” (0.004), where handling the crisis in general. Lastly, Topic 5 covered commuter concerns about “quarantine” (0.008), “thanksgiving” (0.005), and other holidays (0.005) during the pandemic. This topic suggested that members of the public were worried about how they were going to celebrate Thanksgiving in a safe manner.

In the first vaccine dose period, for Topic 1, the word “get” achieved the highest probability of 0.009, followed by “pandemic” (0.008), “people” (0.007), “vaccine” (0.004), and “time” (0.004), implying that the conversation centered around getting people to get vaccinated. The second and third topic still centered on getting the vaccine before using trains and buses. The need for people to get vaccinated is still emphasized in the fourth and fifth topics but includes using masks while on buses and trains to reduce the spread of COVID-19.

According to Table 3, Topics 1 through 4 of the tweets related to the second vaccine dose showed that with more “people” (0.005) taking the “vaccines” (0.006), they can integrate more in the society without the worry of transmitting the virus at “home” (0.007). Fully vaccinated (0.005) people (0.010) (i.e., individuals who have received the first and second dose) gradually returned to taking public transportation to work in different states across the US. Also, some public discourse centered on wearing a face “mask” (0.009) while using public “transit” (0.006) and at “work” (0.004) even if an individual has taken the “vaccine” (0.008), as observed in the fifth topic. Overall, the section shows the public’s commitment to return to their daily lives and routines after staying home for an extended period.

In the Delta variant period, for Topic 1, it is seen that “mask” has the highest probability value of 0.017, followed by “bus” (0.016), “train” (0.009), “people” (0.009), and “wear” (0.008). The new variant caused an increase in infection rates around communities leading to more public discussion and an increased sense of safety when using public transportation. The second and third topics deal with fully “vaccinated” (0.008) “people” (0.007) being able to go to “work” (0.006) using public “transit” (0.004) regardless of the new variant. The fourth topic concerns vaccination in “schools” (0.007), and as a requirement for “international” (0.004) travel. The final topic is centered on the “need” (0.006) for people to be vaccinated (0.006) for protection against the new variant.

According to Table 3, Topic 1 of the Booster Shot and Omicron variant period related tweets have the following clusters: “new” (0.007), “variant” (0.007), “train” (0.005), “vaccinated” (0.005), “back” (0.005), “restrictions” (0.005), “take” (0.004), “ban” (0.004), and “Africa” (0.004). After evaluating the tweets containing these words, it was found that users were concerned about the new variant and were confident in the new vaccines (i.e., the booster shot) but voiced their displeasure at the travel restrictions and bans placed on African countries. They were concerned that restricting travel from African countries was biased because the new variant was found there. Topics 2 and 3 show the public’s stance towards the “Biden” (0.004) administration on removing the travel “restriction” (0.006) and “bans” (0.005) placed on South Africa. The “bans” (0.010) and new “variant” (0.009) were the major discussion points in Topic 4. Finally, Topic 5 centers on people wearing face “masks” (0.006) to protect themselves from the “Omicron” (0.006) variant while using “buses” (0.005) and “trains” (0.005). Also, Twitter users encouraged people to get the new “vaccines” (0.004) for increased protection.

4. Summary & Conclusions

Conclusions

In this chapter, we summarize the research questions that led to this study. Additionally, we identify policy implications of the findings for the current study and acknowledge the study's limitations. Finally, we recommend areas of opportunity for future work.

Summary of Findings

This report concludes by highlighting a list of significant findings, which can be used as a proxy for policy development regarding the public's perception of transportation issues in a pandemic. This conclusion was made based on the sentiment patterns identified and reported in this study and their geolocation.

The pandemic adversely affected public transportation usage as well as related businesses. Many people avoided the use of public transportation and instead preferred private/personal vehicles. In the worst case, people preferred not to go out due to fear of contracting the virus, and overall, this affected businesses that depend on public transportation directly or indirectly. One such situation is the reduction in gas prices and mass layoffs by gas producers. Additionally, service stations were closing and cutting down workers' hours as they recorded low patronage.

There was a high probability of spreading the COVID-19 virus while using public transportation systems. The spread of the virus was a serious concern for commuters as they felt that strict measures like social distancing and the wearing of face masks could not be fully enforced because of diverse beliefs held by commuters about the existence of the virus and its spread majorly due to mixed-messaging. Some people found it difficult to accept that the virus was deadly and different from influenza. Also, it was out of place for many to believe a vaccine is needed. Therefore, they remained careless and refused to follow the suggested health guidelines given by public health experts. These acts made public transportation less attractive for other commuters who believed in the existence of the COVID-19 virus.

Since the vaccination status of public transit users was unknown, most riders could not make informed decisions about how to best protect themselves. Although several studies revealed that face masks help to reduce the spread and transmission of SARS-CoV-2 infection, many users do not want to rely on them when using public transportation.

The infection rates were correlated with public transportation engagements. Increased infection rates within communities caused more discussion about public transportation among Twitter users and vice versa. As more people realized that infection rates were reported from public transport users, people became increasingly worried about their safety, thereby reducing confidence in the use of public transportation, and causing a reduction in patronage. As a result of vaccination and

stricter social distancing measures imposed on the citizenry, we noticed a decline in social media engagement on public transportation, and an increase in the use of public transportation.

At any given time, there were multiple lines of evidence that showed the public expressing their concerns on the use of public transportation when there was an increasing spread or emergence of a new COVID-19 variant. This showed that social media platforms can be a promising medium for understanding the community's perspective of public transport. The confidence of transit users was elevated in the later months of the pandemic. People were adopting the messaging from public health authorities as they were getting vaccinated, and more strict measures were being adopted by both individuals and public transport authorities. This improved public transportation use and reduced the amount of fear being expressed through social media.

Implications for Policymakers Toward a Continued Public Transportation Engagement

As a result of the COVID-19 pandemic, public transportation, and its associated businesses faced difficulties. People in those businesses faced mass layoffs. Gas stations were less frequented, car repair shops went days without patronage, and the economy was crumbling. To fight these issues, with the outcomes of this study, we suggest several strategies that policymakers can develop a pandemic public transportation plan on how to reduce or circumvent the adverse effects faced during a pandemic.

Our first concrete suggestion is to develop a model that continuously harvests social media data, as it contains transport feedback information, that can be used to improve transportation systems. For instance, Twitter users encouraged one another to wear face coverings, maintain social distance, and use sanitizers while traveling in public transit. At different stages of the pandemic, users urged governments and transit agencies to implement advanced technologies, such as Google Maps data, to alert transit riders and operators about COVID-19 hotspots and test centers. Some of these approaches will ensure that public commuters become more adaptable in their adherence to safe commuting measures, thereby reducing the spread of the virus, and increasing public confidence in the continuous use of public transportation.

Alternative channels should be explored, and the right influencers (e.g., trusted public officials) should assist in disseminating policies and guidelines. It was observed that when accurate and clear health information was conveyed to the public, it increased the acceptability of the vaccine and improved the numbers who used public transport systems.

Other mediums of communication could be used to convey guidelines and policies. The use of text messaging with links that show areas of high infection and transmission rates could help guide commuters on how to protect themselves and their loved ones if they need to be in those hotspot areas.

Limitations of the Study and Recommendations for Future Research

We identified multiple lines of evidence on issues surrounding the impact of COVID-19 on public transportation usage using sentiment analysis from Twitter users. First, we acknowledge the relatively small sample size used for this analysis, compared to the entire population of the United States. Second, this study used only tweets from English-speaking users within the United States. The information from these users may not represent the concerns of non-English users. Their views are likely to make the findings of this study more robust. Third, we did not consider the opinions of transport users in the study. Fourth, we did not collect demographic-related information (e.g., age, race, education, or income) from the users. So, we cannot tell whether users fall into under-represented groups. Finally, the scope of this study did not go beyond compact modes of public transport such as cars and air travel. We suggest that future work can investigate how other modes of transportation, such as bicycles, were affected during the pandemic.

Bibliography

- Arafat, M. (2020). A review of models for hydrating large-scale twitter data of covid-19-related tweets for transportation research.
- Astroza, S., Tirachini, A., Hurtubia, R., Carrasco, J. A., Guevara, A., Munizaga, M., . . . Torres, V. (2020). Mobility Changes, Teleworking, and Remote Communication during the COVID-19 Pandemic in Chile. *Findings*, 13489.
- BART. (2020). Ridership Watch: daily updates related to riders returning to BART. Retrieved from <https://www.bart.gov/news/articles/2020/news20200225>
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3(Jan), 993-1022.
- BTS. (2022). Latest Weekly COVID-19 Transportation Statistics. Retrieved from <https://www.bts.gov/covid-19/week-in-transportation>
- CDC. (2022). COVID Data Tracker Weekly Review. Retrieved from <https://www.cdc.gov/coronavirus/2019-ncov/covid-data/covidview/past-reports/01282022.html>
- Das, S., Sun, X., & Dutta, A. (2016). Text mining and topic modeling of compendiums of papers from transportation research board annual meetings. *Transportation research record*, 2552(1), 48-56.
- Gkiotsalitis, K., & Cats, O. (2021). Optimal frequency setting of metro services in the age of COVID-19 distancing measures. *Transportmetrica A: Transport Science*, 1-21.
- Habib, M. A., & Anik, M. A. H. (2021). Impacts of COVID-19 on Transport Modes and Mobility Behavior: Analysis of Public Discourse in Twitter. *Transportation research record*, 03611981211029926.
- Hussein, D. M. E.-D. M. (2018). A survey on sentiment analysis challenges. *Journal of King Saud University-Engineering Sciences*, 30(4), 330-338.
- Liu, B. (2010). Sentiment analysis and subjectivity. *Handbook of natural language processing*, 2(2010), 627-666.
- Liu, L., Tang, L., Dong, W., Yao, S., & Zhou, W. (2016). An overview of topic modeling and its current applications in bioinformatics. *SpringerPlus*, 5(1), 1-22.

- Molloy, J., Schatzmann, T., Schoeman, B., Tchervenkova, C., Hintermann, B., & Axhausen, K. W. (2021). Observed impacts of COVID-19 on travel behaviour in Switzerland based on a large GPS panel. *Transport Policy*.
- Osorio-Arjona, J., Horak, J., Svoboda, R., & García-Ruiz, Y. (2021). Social media semantic perceptions on Madrid Metro system: Using Twitter data to link complaints to space. *Sustainable cities and society*, 64, 102530.
- Taleqani, A. R., Ezekwem, K., Molina, A., Askarzadeh, T., Malalgoda, N., Hough, J., & Mattson, J. (2021). *Using Topic Modeling to Identify Public Opinion on Public Transportation During the COVID-19 Pandemic*. Retrieved from
- Tirachini, A., & Cats, O. (2020). COVID-19 and public transportation: Current assessment, prospects, and research needs. *Journal of Public Transportation*, 22(1), 1.
- TransitApp. (2021). How Coronavirus Is Disrupting Public Transit. Retrieved from <https://transitapp.com/coronavirus>
- Zhang, L., Wang, S., & Liu, B. (2018). Deep learning for sentiment analysis: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(4), e1253.

About the Authors

Egbe-Etu Etu, Ph.D.

Dr. Etu is an Assistant Professor of Business Analytics at San Jose State University (SJSU). He is also a Research Associate in the Mineta Transportation Institute. Before joining SJSU, Dr. Etu received his Ph.D. in Industrial and Systems Engineering from Wayne State University, Detroit in 2021 and his bachelor's degree in Civil Engineering from Covenant University, Nigeria, in 2016. His research interest centers on the development of use-inspired machine learning models to solve challenging business problems in healthcare, manufacturing, and transportation. He is a member of the Industrial Engineering and Operations Management (IEOM), the Institute of Industrial & Systems Engineering (IISE), and SAVE International.

Imokhai Tenebe, Ph.D.

Dr. Tenebe is currently a software engineer. Previously, he was a safety engineer with the Texas Commission on Environmental Quality in Austin, Texas and participated in research at Texas State University, San Marcos. He obtained his Ph.D. in Water Resources and Environmental Engineering and has two masters in the same area of expertise from Nigeria and the United States. He has published over 80 articles on several subjects with interest in water resources, healthcare, pollution, transportation, and data science. Dr. Tenebe is also a Research Associate at the Mineta Transportation Institute.

Ankur Parma

Ankur is a graduate student completing his M.S. in artificial intelligence at San Jose State University. He received his B.S. in Applied Petroleum Engineering from the University of Petroleum and Energy Studies, India, in 2016. His research interests lie in deep learning and reinforcement learning.

Likhitha Yelamanchili

Likhitha is a graduate student completing her M.S. in computer science at San Jose State University. She received her Bachelor of Technology in computer science from the Gayatri Vidya Parishad College of Engineering (Autonomous), India, in 2019. Her research interests include machine learning, big data, and artificial intelligence. Likhitha also enjoys developing web applications.

Dang Minh Nhu Nguyen

Nguyen is a recent graduate of San Jose State University. Nguyen received her B.S. in Applied Mathematics, with a concentration in Statistics in June 2022. Her research interests include technology, healthcare, business, finance, and aerospace.

Louis Tran

Louis is an undergraduate student majoring in Applied Mathematics, with a concentration in Statistics at San Jose State University. His research interests lie in machine learning methods and applications.

Ihor Markevych

Ihor received his M.Sc. in computational data science from Carnegie Mellon University, Pittsburgh in 2021 and his bachelor's degree at the Institute of Applied System Analysis from the National Technical University of Ukraine in 2019. His research interests include differential modeling of semi-linear parabolic systems with composition methods and convergence of iterations in the Trotter-Daletskii formula for nonlinear perturbation. He is an active member of the Association of Computing Machinery Society.

Hon. Norman Y. Mineta

MTI BOARD OF TRUSTEES

Founder, Honorable Norman Mineta***
Secretary (ret.),
US Department of Transportation

**Chair,
Will Kempton**
Retired Transportation Executive

**Vice Chair,
Jeff Morales**
Managing Principal
InfraStrategies, LLC

**Executive Director,
Karen Philbrick, PhD***
Mineta Transportation Institute
San José State University

Winsome Bowen
President
Authentic Execution Corporation

David Castagnetti
Co-Founder
Mehlman Castagnetti
Rosen & Thomas

Maria Cino
Vice President
America & U.S. Government
Relations Hewlett-Packard Enterprise

Grace Crunican**
Owner
Crunican LLC

Donna DeMartino
Retired Transportation Executive

John Flaherty
Senior Fellow
Silicon Valley American
Leadership Forum

Stephen J. Gardner*
President & CEO
Amtrak

Rose Guilbault
Board Member
San Mateo County
Transit District (SamTrans)

Kyle Christina Holland
Senior Director,
Special Projects, TAP Technologies,
Los Angeles County Metropolitan
Transportation Authority (LA Metro)

Ian Jefferies*
President & CEO
Association of American Railroads

Diane Woodend Jones
Principal & Chair of Board
Lea + Elliott, Inc.

Therese McMillan
Executive Director
Metropolitan Transportation
Commission (MTC)

Abbas Mohaddes
CEO
Econolite Group Inc.

Stephen Morrissey
Vice President – Regulatory and
Policy
United Airlines

Toks Omishakin*
Secretary
California State Transportation
Agency (CALSTA)

Takayoshi (Taki) Oshima
Chairman & CEO
Allied Telesis, Inc.

Marco Pagani, PhD*
Interim Dean
Lucas College and
Graduate School of Business
San José State University

April Rai
President & CEO
Conference of Minority
Transportation Officials (COMTO)

Greg Regan*
President
Transportation Trades Department,
AFL-CIO

Paul Skoutelas*
President & CEO
American Public Transportation
Association (APTA)

Kimberly Slaughter
CEO
Systra USA

Tony Tavares*
Director
California Department of
Transportation (Caltrans)

Jim Tymon*
Executive Director
American Association of
State Highway and Transportation
Officials (AASHTO)

* = Ex-Officio
** = Past Chair, Board of Trustees
*** = Deceased

Directors

Karen Philbrick, PhD
Executive Director

Hilary Nixon, PhD
Deputy Executive Director

Asha Weinstein Agrawal, PhD
Education Director
National Transportation Finance
Center Director

Brian Michael Jenkins
National Transportation Security
Center Director

