

Global Prevalence Patterns of Anti-Asian Prejudice on Twitter During the COVID-19 Pandemic

Brittany Wheeler, Monika Purohit, Patrick Furman, Seong Jung, Deborah L. Hall, Maria Camila N. Barioni, and Yasin N. Silva

Abstract Anti-Asian prejudice increased during the COVID-19 pandemic, evidenced by a rise in physical attacks on individuals of Asian descent. Concurrently, as many governments enacted stay-at-home mandates, the spread of anti-Asian content increased in online spaces, including social media platforms such as Twitter. In the present study, we investigated temporal and geographic patterns in the prevalence of social media content relevant to anti-Asian prejudice within the U.S. and worldwide. Specifically, we used the Twitter Data Collection API to query over 13 million tweets posted during the first 15 months of the pandemic (i.e., from January 30, 2020 to April 30, 2021), for both negative (e.g., #kungflu) and positive (e.g., #stopAAPI-hate) hashtags and keywords related to anti-Asian prejudice. Results of a range of exploratory and descriptive analyses offer novel insights. For instance, in the U.S., results from a burst analysis indicated that the prevalence of negative (anti-Asian) and positive (counter-hate) messages fluctuated over time in patterns that largely mirrored salient events relevant to COVID-19 (e.g., political tweets, highly-visible hate crimes targeting Asians). Other representative findings include geographic differences in the frequency of negative and positive keywords that shed light on the regions within the U.S. and the countries worldwide in which negative and positive messages were most frequent. Additional analyses revealed informative patterns in the prevalence of original tweets versus retweets, the co-occurrence of negative and positive content within a tweet, and fluctuations in content in relation to the number of new COVID-19 cases and reported COVID-related deaths. Together, these findings underscore the value of research examining trends in social media messages of hate and counter-hate during the COVID-19 pandemic.

Brittany Wheeler, Seong Jung, Deborah L. Hall
Arizona State University, Glendale, Arizona, e-mail: [bmwheel5,shjung3,d.hall]@asu.edu

Monika Purohit, Patrick Furman, Yasin N. Silva
Loyola University Chicago, Chicago, Illinois e-mail: [mpurohit,pfurman,ysilva1]@luc.edu

Maria Camila N. Barioni
Federal University of Uberlândia, Uberlândia, Brazil e-mail: camila.barioni@ufu.br

1 Introduction

The World Health Organization (WHO) reported on January 4, 2020, that it was monitoring an outbreak of a new virus in the Wuhan, Hubei Province of China [1]. At this time, knowledge of and concern about the virus from the public was limited. Less than one month later, however, on January 30, the WHO declared the spread of the virus, termed COVID-19, a public health emergency, bringing global attention to this widespread health concern [1, 2]. The name ‘coronavirus’ was developed according to WHO’s “Best Practices for the Naming of New Human Infectious Diseases,” which recommends avoiding cultural, social, regional, or ethnic associations when naming a disease [3]. Despite these recommendations, because of the origins of the virus, COVID-19 was frequently referred to in the media as the “Chinese virus,” the “Wuhan virus,” and the “Asian virus” [4, 5, 6, 7]. To illustrate, a widely-shared image revealed that President Trump had crossed out the word “corona” in “coronavirus” from his notes and replaced it with the word “Chinese” [8].

While some have argued that this terminology is not inherently racist, given the virus’ origin, anti-Asian prejudice notably increased in prevalence and salience during this time [9, 10, 11, 12, 13]. For example, police reports in the U.S. involving anti-Asian hate and physical violence against Asian Americans and Pacific Islanders (AAPI) increased 145% in 2020 compared to previous years [11] and Stop AAPI Hate—a non-profit organization dedicated to reducing anti-Asian prejudice—reported 2,583 incidents of anti-Asian prejudice between March 18, 2020 and August 5, 2020 [15]. This number of incidents may, in fact, be higher given that these only represent reported instances and do not account for unreported acts of hate or harassment [16].

Online anti-Asian hate speech, defined as any online message that includes “profanity, offensive language, or toxicity” directed at Asian individuals [18], also increased during this time. The Anti-Defamation League, for instance, reported an 85% increase in anti-Asian discrimination online [19]. To illustrate, during the first months of the pandemic, 72,000 posts on Instagram contained the hashtag #Wuhan-Virus, while another 10,000 contained the hashtag #KungFlu [20]. Notably, social media posts (i.e., tweets) generated by President Trump and other political leaders used the phrase “Chinese Virus” [9]. The role of these tweets in promoting the continued use of the term is perhaps reflected by the finding that 18% of tweets using anti-Asian hashtags referred to Trump in some capacity [9, 21]. In fact, recent research by Kim and Kesari [22] identified marked increases in anti-Asian terminology after President Trump first started using similar language.

Whereas some have argued that the use of this language is benign, others have found its deliberate and continued use signals support for anti-Asian attitudes and contributes to the proliferation and wider dissemination of hateful messages [9, 23]. The increased prevalence of hate in online spaces is also particularly concerning, given that it has paralleled increases in real-life (i.e., offline) hate crimes [17, 24]. Indeed, the use of these terms is one of the most frequently cited reasons for the increase in anti-Asian hate crimes, according to individuals of Asian descent [10, 11].

Research has demonstrated that, despite its widespread nature, hateful content on Twitter tends to be produced by a small number of users—an alarming finding given the amount of content produced [25]. Interestingly, counter-hate content (i.e., positive language intended to combat hateful messages) that drew connections between anti-Asian terminology and xenophobia and prejudice also increased during this time [22]. Producers of counter-hate content were more likely to be women, younger adults, and—perhaps not surprisingly in light of the profound impact on their lives—Asian-American reporters, professors, and politicians [26].

Although a number of studies on anti-Asian sentiment during the COVID-19 pandemic have been conducted within the United States, comparatively fewer studies have taken a global perspective on this issue. Yet, it is important to recognize that violent attacks and hate speech online are not limited to one particular geographic location. Reports from the United Kingdom, for instance, have shown that Asian individuals were also subjected to physical assaults and exclusion from various social settings [27, 28, 29].

In order to better understand national-level cultural differences related to anti-Asian sentiment, Hofstede's cultural dimensions [30, 31] have been used in previous research. Hofstede's framework suggests that there are several key factors associated with evaluating cultural differences, including individualism (i.e., the extent to which individuals are integrated within the community), masculinity (i.e., the extent to which emotional roles are distributed evenly across genders), uncertainty avoidance (i.e., the extent to which individuals within a culture can feel comfortable in uncertain situations), and power distance (i.e., the extent to which power is distributed equally across members of the community).

Building on this, Ng [31] found that cultures high in individualism, power distance, and uncertainty avoidance, such as the United Kingdom, India, and Ghana, exhibited higher levels of negative, anti-Asian sentiment during the pandemic. This phenomenon can be attributed to various factors that are unique to each cultural dimension. In societies that value individualism, individuals tend to prioritize their personal opinions and express their emotions more openly [31]. Consequently, there was a greater tendency to direct blame and anger toward Asian individuals during the pandemic, resulting in a higher level of negative sentiments [31]. Cultures with high power distance tend to exhibit greater inequality, which perpetuates social dichotomies and fuels xenophobic sentiments toward individuals perceived to have a lower status. Thus, high power distance predicted a heightened level of negative sentiment directed toward Asian individuals, who were deemed unworthy of respect due to their perceived lower status [31]. Cultures with high uncertainty avoidance tend to label others based on a variety of stigmas that reinforce negative stereotypes and perpetuate a lack of tolerance toward certain groups of people [31]. In the case of the COVID-19 pandemic, this construct was also associated with increased negative sentiment directed toward Asian individuals [31].

In addition to research that adopts a global perspective, studies that build on and speak to the relationship between online hate and real-world outcomes [17, 24] may be particularly beneficial. By the middle of 2021, the COVID-19 pandemic had caused the loss of over four million lives globally, with more than 188 mil-

lion individuals contracting the virus [32]. One important consideration is how the prevalence of online anti-Asian content may have varied in relation to COVID-19 milestones and transmission rates, such as the number of new COVID-19 cases and COVID-19-related deaths.

Finally, it is worth noting that prejudicial attitudes have also been observed during other infectious disease outbreaks [33, 34, 35, 36]. In the U.S., the 2009 H1N1 outbreak originating in Mexico led to a surge in stigma against Mexicans and people of Hispanic descent [34, 36]. Similarly, after the 2003 Severe Acute Respiratory Syndrome (SARS) outbreak, which was linked to China, Asian Americans were targeted with disease-related stigma [33]. A crucial difference between these prior outbreaks and the COVID-19 pandemic, however, is the pervasiveness of social media in the daily lives of individuals on a global scale. Thus, understanding temporal and geographic patterns in anti-Asian content online during the COVID-19 pandemic—especially in the context of influential events and pandemic-related milestones—can provide unique insights about the spread of online hate, more generally, during global public health crises.

In this paper, we present the results of a range of exploratory and descriptive analyses to shed light on anti-Asian prejudice—as well as messages to counter this hate—on Twitter during a 15-month span of the COVID-19 pandemic. We begin by reviewing findings originally discussed in our previous work [37] that offer insights on temporal and geographic patterns in the frequency of anti-Asian and counter-hate content. Crucially, however, we extend our previous work in several important ways. For instance, one concern with the frequent use of anti-Asian rhetoric on social media has been the potential for the re-circulation of this content. On Twitter, re-circulation can occur through replies (i.e., a tweet directed at the original poster by utilizing the ‘reply’ button) and retweets (i.e., a pure echoing of content created by another user) [38]. Retweeting hateful anti-Asian content is particularly harmful, as it allows the content to be amplified and reach new and wider audiences, signals support from a listening audience, and validates the original poster’s ideas [39]. While limited research has evaluated this re-circulation with a focus on anti-Asian messages, research by Lindgren [38] examining other online movements found that increased retweets can be viewed as micro-acts of support. This is further illustrated by the work of Solovev and Prollochs [40], who found that tweets initially containing more negative sentiment are more likely to contain higher levels of hate speech within replies.

In our previous work [37], we sought to contribute to the body of research regarding anti-Asian prejudice by covering a significantly longer time frame, considering both anti-Asian and positive keywords, and employing a relatively under-utilized method (i.e., burst analysis) to understand trends in keyword use associated with anti-Asian prejudice. In the present paper, we seek to expand on these findings by taking a more fine-grained approach to evaluating anti-Asian and counter-hate content on Twitter. Specifically, in our previous work, we covered broad trends in the use of specific hateful and counter-hate terminology—which, although beneficial, does not account for the differences in how users can engage with this content (i.e., reply, retweet, post original content). Additionally, given that previous research has indi-

cated that most hateful content is produced by a relatively limited number of users [25], we seek to evaluate trends in anti-Asian and counter-hate messages within the top 10% of producers of relevant content. That is, understanding the trends associated with the most “influential” users can provide valuable insights, given their centrality within online discussions. Also, while global trends were touched upon within our previous paper, we dis-aggregate this content to better understand negative and positive content trends from the top three countries producing this content. Drawing on Hofstede’s work on cultural dimensions [30], we aim to develop an understanding of why certain countries generated more anti-Asian and counter-hate content. Furthermore, new additions to this work include an evaluation of the co-occurrence of anti-Asian and counter-hate keywords at the level of individual tweets and a descriptive analysis of reply conversion (i.e., when replies to a message change sentiment). These analyses can provide a better understanding of the interplay of anti-Asian and counter-hate messages within tweets and the subsequent online interactions they elicit. Lastly, we re-generated several of the figures presented in our previous work which only considered U.S. data with a broader focus on global trends and integrated data on new COVID-19 cases and COVID-related deaths with temporal trends in anti-Asian and counter-hate tweets.

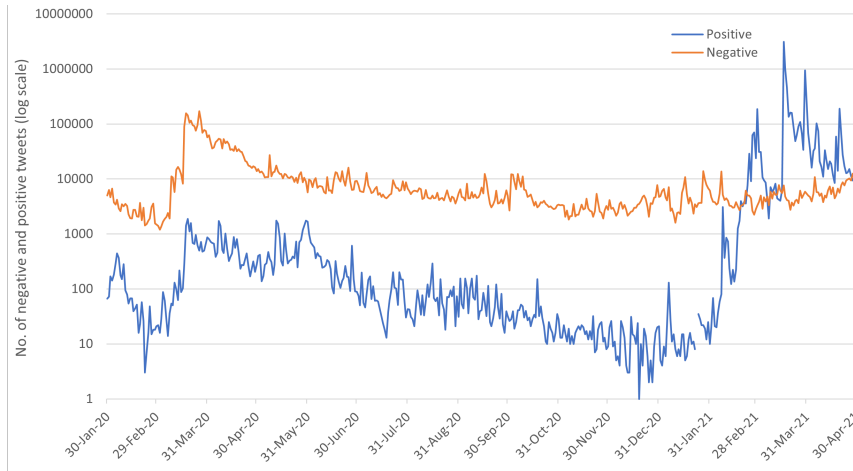
The remainder of the paper is organized as follows: In Section II, we discuss in greater detail the procedures we employed to collect and clean two complementary sets of Twitter data. In Section III, we present the results of a set of diverse analyses examining temporal and geographic patterns in the data, as well as exploratory findings that shed light on more nuanced aspects of anti-Asian and counter-hate content. Finally, in Section IV, we discuss implications, limitations, and crucial directions for future research.

2 Methods

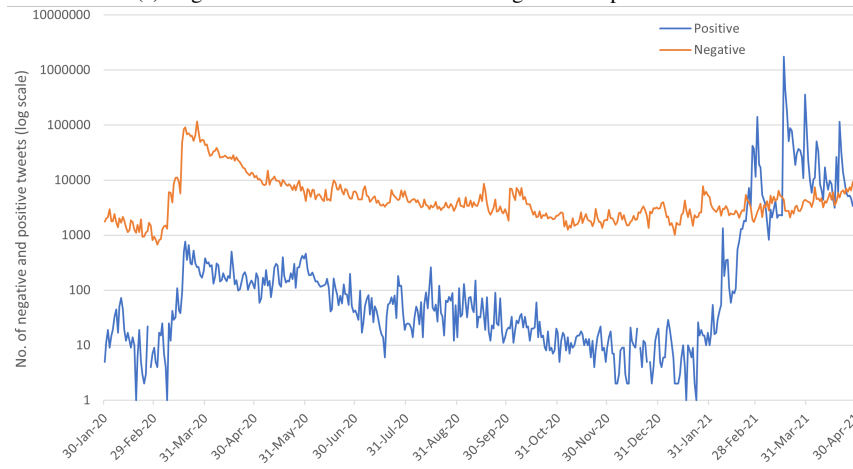
2.1 Data

All data were collected according to Twitter data collection guidelines and using the proper API access provided to researchers [41, 42]. In the following sections, we refer to anti-Asian content as “negative” and counter-hate content as “positive.”

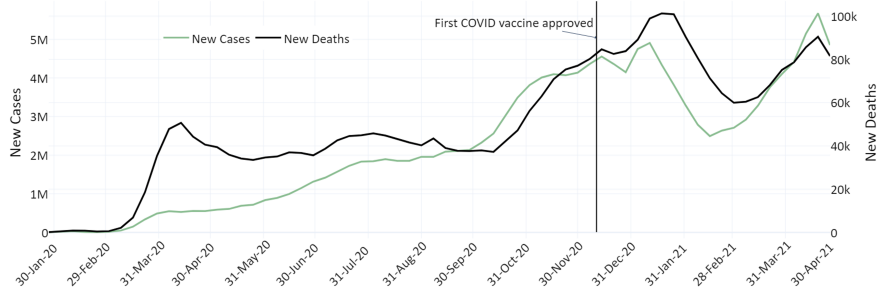
Archive Dataset. Using the Twitter Data Collection API [43], we queried tweets containing negative and positive hashtags and keywords related to anti-Asian prejudice from January 30, 2020 to April 30, 2021. This time frame was selected to correspond with the date on which the World Health Organization indicated the spread of COVID-19 was a global health issue and the start of AAPI Heritage Month the following year (when positive AAPI messages might increase independent of COVID-19). We used 12 specific negative hashtags/keywords as indicators of anti-Asian prejudice (#batsoup, #chinavirus, #gobacktochina, #chinesevirus, #chineseplague, #gook, #chinaliedpeople, #kungflu, #wuflu, #chingchong, #makechinapay, #ccpvirus) and 5 specific positive hashtags/keywords (#hateisavirus, #iamnotavirus, #racis-



(a) Logarithmic scale of the number of negative and positive tweets.



(b) Logarithmic scale of the number of negative and positive tweets for the top 10% of producers.



(c) Number of new COVID-19 cases and deaths by week.

Fig. 1 Global number of negative and positive tweets and the number of new COVID-19 cases and deaths between January 2020 and April 2021.

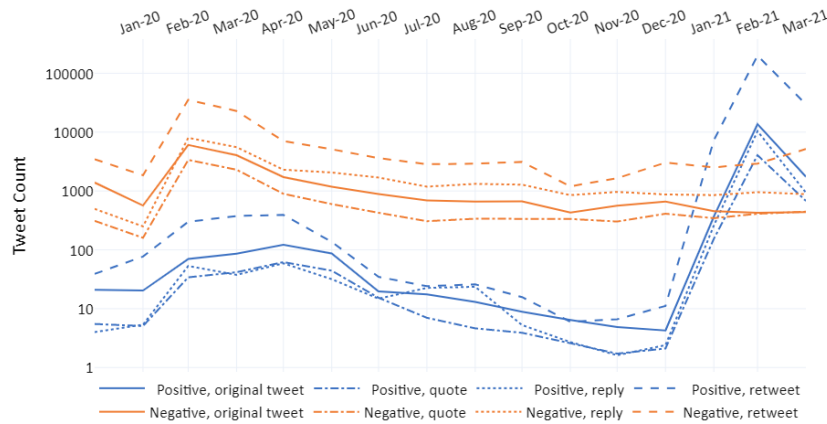


Fig. 2 Number of negative and positive tweets between January 2020 and April 2021 by type of tweet.

misavirus, #washthehate, #stopasianhate), which were chosen based on the relevant literature [44], news publications [45, 46, 47], and social media posts discussing anti-Asian attitudes during the beginning of the pandemic. The total sample consisted of 13,008,053 tweets from 3,298,940 distinct users.

1% Dataset. The 1% sample stream dataset was generated from Twitter’s sample stream endpoint [49] which provides access to a roughly 1% random sample of publicly available tweets in real-time. This dataset was compiled from the tweets gathered over the course of 24 hours (August 1-2, 2021) to estimate the amount of activity that 1% of the Twitter platform could generate in one day. 4,093,933 tweets were collected in this sample from 2,956,806 distinct users.

2.2 Geographic Location Labeling.

During the collection of both datasets, a filter was applied to collect a list of users who had publicly available geolocation data through their location settings. To perform descriptive analyses based on geographic location, we devised a geolocation labeling strategy similar to Jiang and colleagues [48]. This strategy was necessary because less than 0.5% of the tweets in our dataset had available `geo_place` information. For our analysis, we considered the state granularity for the tweets originating in the U.S. and the country granularity for the tweets originating in other countries, based on self-reported user profile locations. Using a fuzzy text matching algorithm [50], pre-processed user-reported locations were matched against a set of predetermined locations inside and outside of the U.S. The similarity between user-reported locations and predetermined locations was computed using the edit distance metric. The

score threshold to consider a matching pair of locations, which was set to 80%, was defined based on a validation analysis conducted by an external human annotator who manually verified a random sample of labeled locations considering country names and U.S. state names. Considering the precision validation measure ($TP/(TP + FP)$), the geolocation labeling strategy achieved a predictive positive value of 99.8% for the U.S. locations and varied from 89.8% to 100% for the other countries that, together with the U.S., account for 90% of the collected data. The set of predetermined locations inside the U.S. consisted of state names and state abbreviations. The set of predetermined locations outside the U.S. was built using the top 20 countries with the most Twitter users as of July 2020 [53] and their 5 most populous cities [54]. To avoid ambiguity, only the country abbreviations that didn't overlap with a U.S. state abbreviation were included. Additionally, we used an ambiguous locations list—built throughout the testing process—to adjust the geolocation labeling by removing ambiguous matches. (An example of an ambiguous match is the token 'valencia,' which can refer to a city in Spain and a town in California.)

2.3 Analytic Plan

Using the previously described datasets, a series of descriptive analyses were performed to investigate the overall prevalence of negative and positive keyword use, as well as trends over time and across locations. We then performed a burst analysis to examine whether there were significant increases in the frequency of negative and positive keywords and, if so, whether such bursts in activity followed similar trends associated with offline hate directed at individuals of Asian descent. (Additional information about the burst detection analysis is presented in the following section.)

3 Results

3.1 Descriptive Analyses

Overall Tweet Frequency. As shown in Fig. 1(a), the use of negative keywords before March 2020 was relatively low. However, within that month, there was a marked increase in the use of negative keywords, with the frequency of negative keywords reaching its peak later in that month. Although considerably less frequent throughout most portions of the timeline, the use of positive keywords, in contrast, culminated in major spikes in late February 2021. Similar trends are also apparent in Fig. 1(b) for the top 10% of users, as the frequency of these negative and positive keywords shows similar peaks in activity. The top 10% producers in our dataset generated 56.1% of tweets and the top 1% were responsible for 24.2%. Fig. 1(c) displays the frequency of new COVID-19 cases and COVID-related deaths globally, based on data from the Centers for Disease Control [55]. Interestingly, there were two

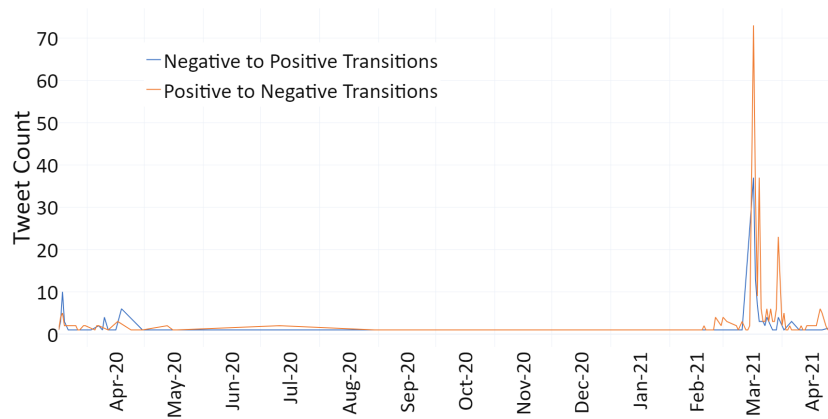


Fig. 3 Frequency of Reply Conversions.

major spikes in the number of reported COVID-related deaths - the first occurring between the months of March 2020 to May 2020, coinciding with the surge in positive keyword frequencies seen in Fig. 1(a). The second spike occurred in early 2021, when the number of positive keywords, especially “stopasianhate,” increased sharply again. These findings suggest a potential link between real-world outcomes, such as the prevalence of COVID-19 cases and deaths, and the generation of counter-hate messages.

Trends by Type of Tweet. To analyze the relation between different types of tweets, we broke down the data presented in Fig. 1(a) to examine trends over time based on tweet type (i.e., original tweet, quote, reply, retweet). Retweets, in particular, may be useful for gauging agreement with an original tweet, as they do not alter its meaning. As depicted in Fig. 2, the vast majority of tweets (78.39%) were retweets, indicating that most tweets containing our selected keywords were not original content. Additionally, while 87.72% of positive tweets were retweets, only 62.99% of negative tweets were retweets, suggesting that tweets containing negative keywords may have been less popular to share overall. Quotes were the least common tweet type, accounting for only 3.75% of tweets. Similar to Fig. 1(a), the dis-aggregated trends in Fig. 2 exhibited comparable patterns over time, with the ratio between the different tweet types remaining relatively constant, particularly for negative tweets. However, we noted a significant spike in the frequency of retweets containing positive keywords around December 2020, compared to subsequent months. This shift was remarkable given that, for several months prior to December, the number of positive retweets was not substantially greater than replies or original tweets. During these same months, negative retweets were far more prevalent than other types of negative tweets, but this dynamic abruptly reversed between December 2020 and February 2021. Replies, unlike retweets, were primarily negative, with 71.76% of them containing negative

keywords, and they were also less common, accounting for only 9.73% of tweets in our dataset.

Sentiment Transitions. Fig. 3 displays daily counts of the number of replies where the original tweet contained a positive keyword and the reply contained a negative keyword or vice versa. We refer to a negative response to a positive tweet as a “positive to negative transition” in the figure. Across all tweets in our data, we only observed 454 transitions of any kind, which amounts to less than 0.04% of the total number of replies—in short, transitions were a particularly rare occurrence. The vast majority of these transitions occurred between February and April of 2021, the same period in which the “stopasianhate” keyword peaked in frequency. Most of these transitions were positive to negative transitions, suggesting a strong response by users of negative keywords to the sudden surge in positive keyword frequency, which was not observed in other time periods.

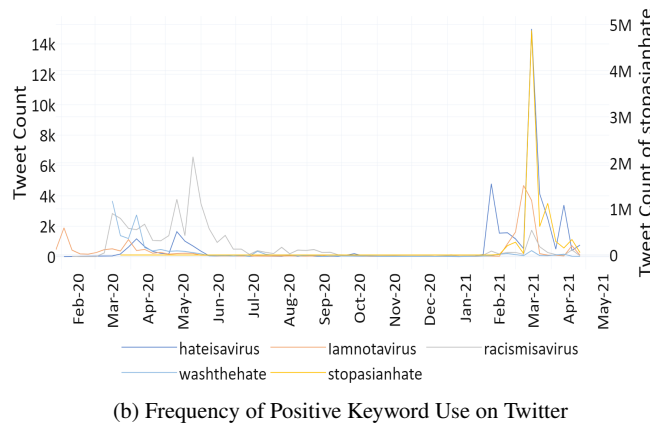
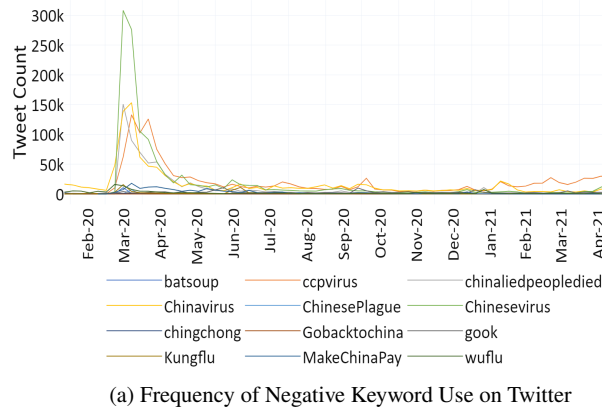


Fig. 4 Frequency of negative and positive keywords on Twitter between January 2020 and April 2021.

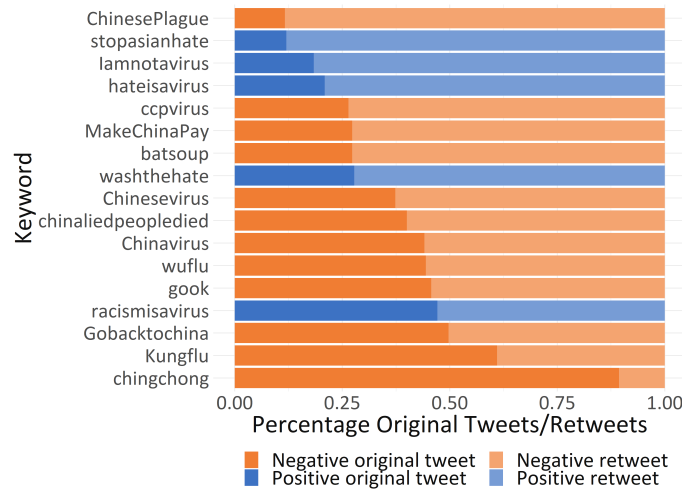


Fig. 5 Percentages of original tweets and retweets for all keywords

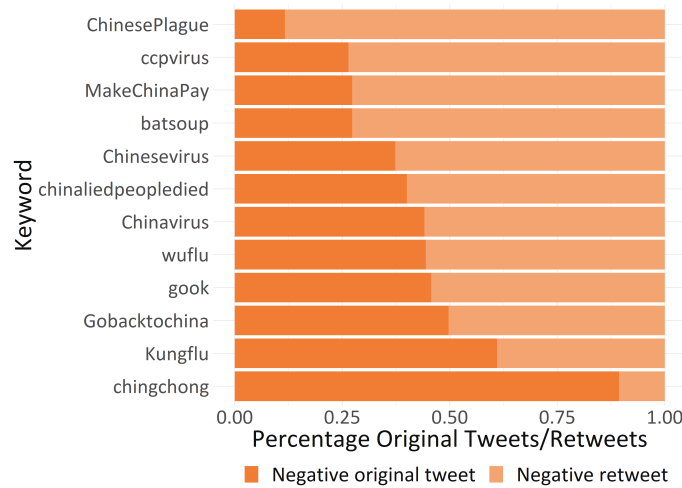


Fig. 6 Percentages of original tweets and retweets for negative keywords

Trends by Type of Keyword. Fig. 4 depicts trends in negative and positive keyword use for the whole dataset over time. As shown, sharp spikes in negative activity occurred following President Trump’s first use of the term “Chinese Virus” in March 2020 [4]. Additionally, although there were more tweets containing positive keywords, overall, these tweets were mainly generated between February 2021 and April 2021. That is, they coincided with a prominent event that occurred in the U.S.—the Atlanta-area spa shootings that resulted in the deaths of multiple individuals of Asian-descent

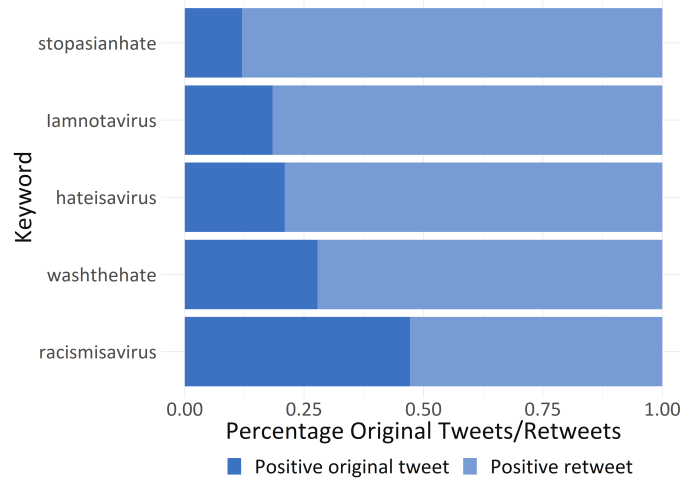


Fig. 7 Percentages of original tweets and retweets for positive keywords

[51]. There was minimal use of these keywords during the early months of the pandemic. As shown in Fig. 4(a), the most frequently used negative keyword was “ccpvirus,” followed by “chinavirus” and “chinesevirus.” The most frequently used positive keywords were “stopasianhate” and “hateisavirus.”

Tweets Vs. Retweets for Individual Keywords. Given the proportion of retweets within the data, we also evaluated the proportion of frequently used keywords within types of retweets, specifically (see Fig. 5, 6, and 7). Three of the four most frequently retweeted keywords were positive, implying that tweets containing negative keywords may have been less appealing to share (i.e., retweet). Furthermore, reflecting its popularity, 88.02% of tweets referencing the “stopasianhate” keyword were retweets. The only keyword with a higher proportion of retweets was “chineseplague,” with 89.35% of tweets being retweets.

Co-occurrence Patterns. Fig. 8 displays the frequency of pairs of keywords used together in the same conversation (defined as an original tweet and all replies to that tweet). We can observe that, except for “chineseplague,” “hateisavirus,” and “racismisavirus,” when more than one keyword appeared in the same conversation, the subsequent keyword(s) tended to be the same (instead of a different) keyword. We did not find this pattern of mirroring, or repetition, for the positive keywords, perhaps because “stopasianhate” alone comprised more than 98% of all positive keyword use. That is, other positive keywords tended to appear alongside “stopasianhate.” One possible explanation is that “iamnotavirus” and “washthehate” were either popular with different users from those who used “stopasianhate” or were popular at different points in time. While some keywords, such as “gook,” were almost never used together with other keywords in a conversation, most keywords showed some level of co-occurrence with other keywords. In cases of co-occurrence, positive keywords

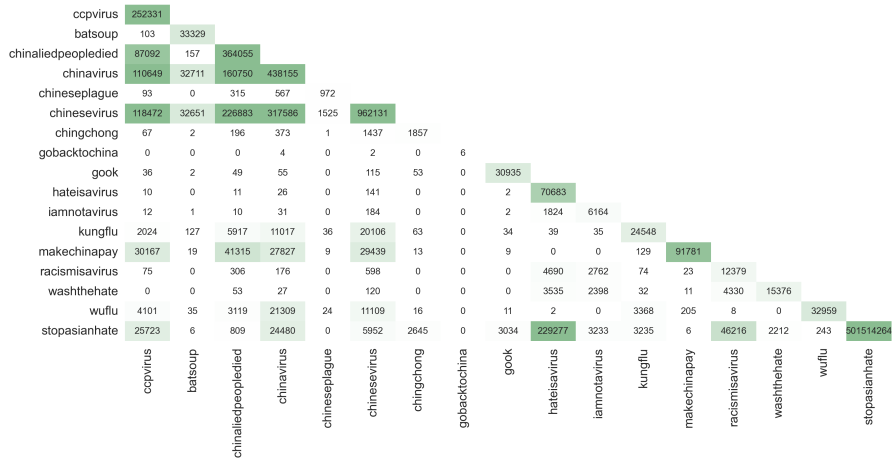
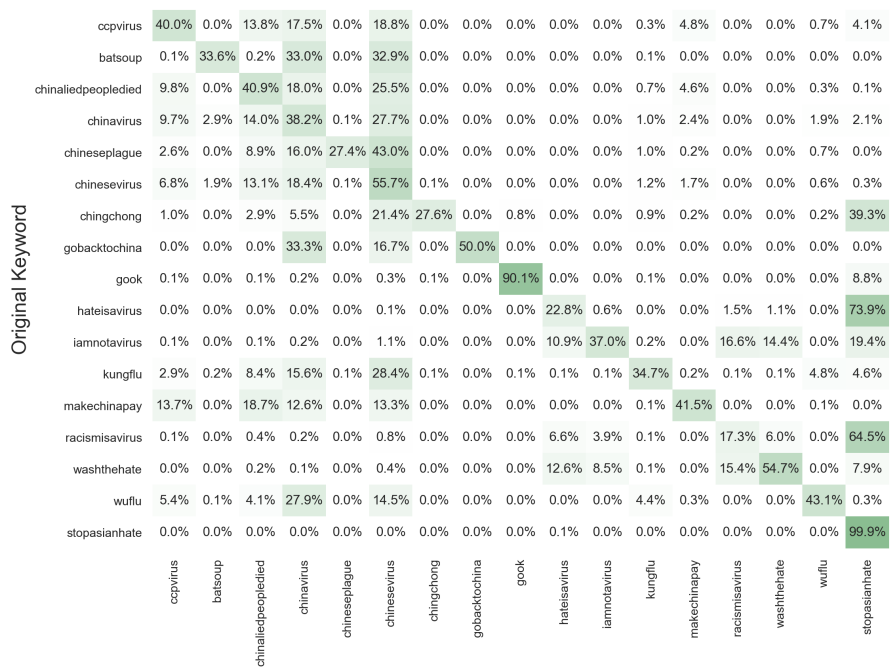


Fig. 8 Frequency table of the number of times keywords co-occur in the same conversation.



Percentage of all keywords used in the same conversation as the original keyword

Fig. 9 Percentage of times keywords co-occur in the same conversation.

tended to be paired with other positive keywords, and negative keywords with other negative keywords.

Notably, the negative keywords “chinesevirus,” “chinavirus,” “chinaliedpeopledied,” “batsoup,” and “ccpvirus” frequently appeared together relative to their individual usage. This implies that these particular negative keywords might have either a similar meaning or a common group of supporters and users. Fig. 9 provides a simplified comparison between keywords of varying frequencies. It is a scaled version of Fig. 8, where each entry represents the percentage of co-occurrences of the original keyword on the left with the corresponding keyword on the bottom, within the same conversation. This is based on the condition that the conversation contained at least two keywords. Although positive keywords rarely appeared with negative keywords and vice versa, “chingchong” was an exception. Specifically, given an initial tweet containing the keyword “chingchong,” 39.3% of keywords that appeared in the same conversation were “stopasianhate.” This may be partly explained by the overwhelmingly greater frequency of “stopasianhate” than “chingchong” and the finding that other negative keywords appeared alongside a positive keyword less than 10% of the time. This suggests that “chingchong” was more likely than other negative keywords to spark discourse on Twitter between users of positive and negative keywords, although further analysis is needed to determine the exact nature of these interactions.

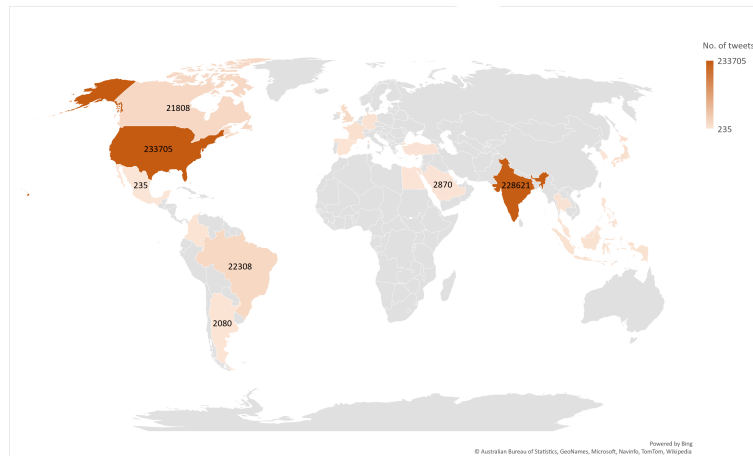


Fig. 10 Count of Negative Tweets per Country

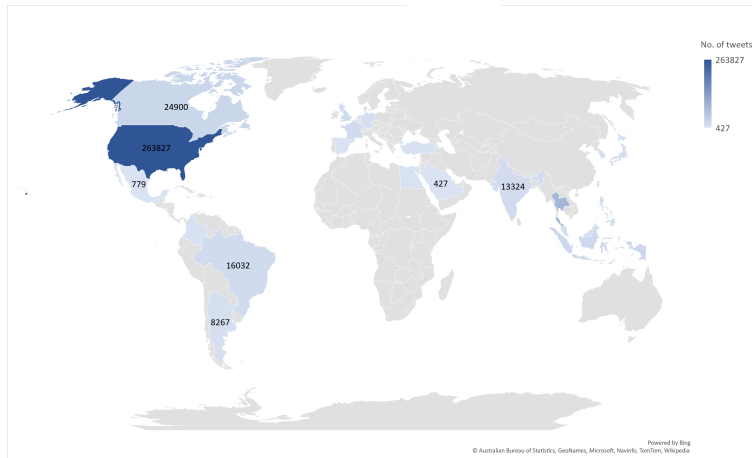


Fig. 11 Count of Positive Tweets per Country

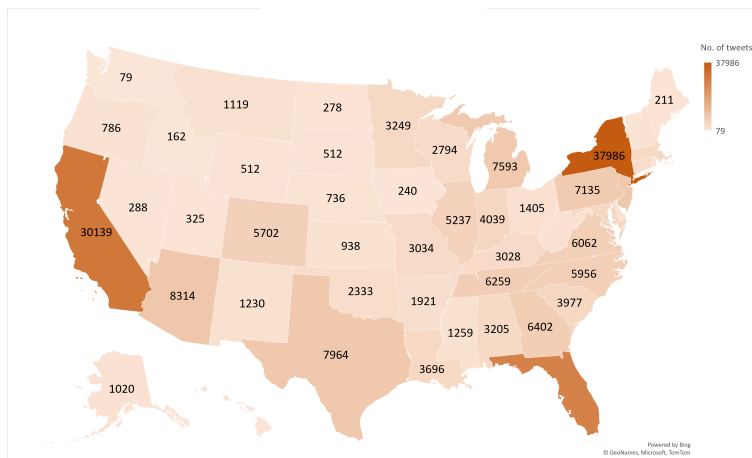


Fig. 12 Count of Negative Tweets in the U.S.

3.2 Regional Trends

Despite its global impact, the COVID-19 pandemic affected different parts of the world at differing rates. Whereas trends in various geographic regions share similarities with the aggregated global trends, exploration of regional differences in negative and positive tweets can yield novel insights.

Negative Keyword Use. As shown in Figures 10-13, globally, 4,521,457 distinct tweets contained at least one of the 12 negative keywords, with most of this content generated in the U.S. and India (USA = 233,705 tweets; IND = 228,621 tweets).

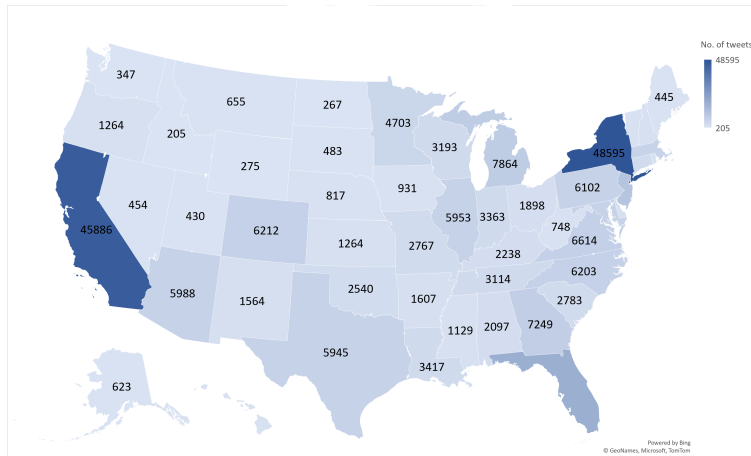


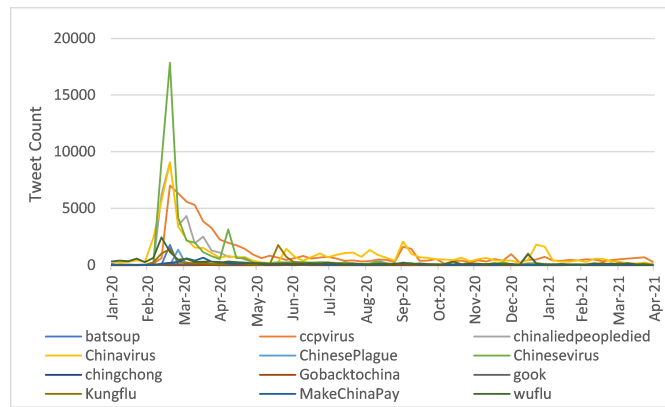
Fig. 13 Count of Positive Tweets in the U.S.

Within the U.S., New York, California, and Florida were the largest producers of negative content (NY = 37,986 tweets; CA = 30,139 tweets; FL = 27,076).

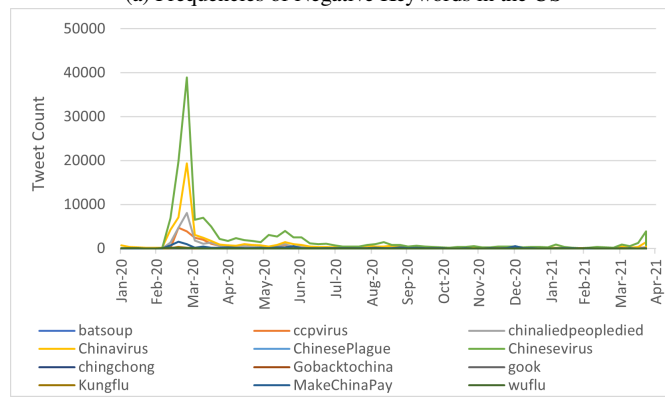
Fig. 14(a) shows that within the U.S., the country with the highest frequency of negative content, most negative keywords were used infrequently until February 2020, after which point, there was a gradual increase. The majority of negative keywords were used between February and March 2020; after this, there was a clear decline in negative keyword use. The term “Chinesevirus” was the most frequently used negative keyword in the U.S., peaking in March 2020 before declining. In March and April 2020, the most commonly used negative keywords were “Chinavirus,” “ccpvirus,” “wufu,” “kungflu,” “makechinapay,” and “Chinesevirus.” Although negative keyword use declined after April 2020, small spikes in “kungflu” and “Chinavirus” were observed in June 2020.

As shown in Fig. 14(b), India—the second-highest producer of negative content—showed similar trends. Negative keyword use in India began in February 2020, peaked between February and July 2020, and then gradually declined. “Chinavirus” was the most frequently used negative keyword in India, mainly in March, April, and July 2020. Other frequently used negative keywords in India were “Chinaliedpeopledied,” “ccpvirus,” and “MakeChinaPay.” Fig. 14(c) shows that in Brazil, the third-highest producer of negative content, negative keyword use did not begin until February 2020, at which point it began to rise. “Chinavirus” was the most frequently used negative keyword in Brazil, peaking in March 2020 before declining. After April 2020, the use of negative keywords in Brazil was minimal.

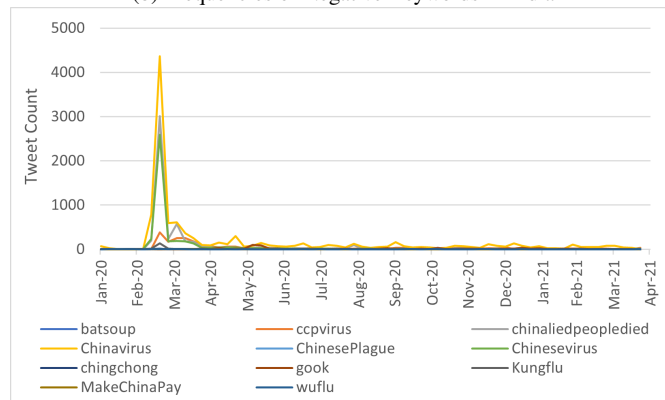
Positive Keyword Use. The analysis of positive content in different regions revealed interesting trends across different countries (Figures 10-13). 6,660,469 distinct tweets contained at least one of the 5 positive keywords, with most of the positive content also generated in the U.S., followed by Thailand (USA = 263,827 tweets; TH = 82,696 tweets). Notably, however, California and New York also produced the most content



(a) Frequencies of Negative Keywords in the US

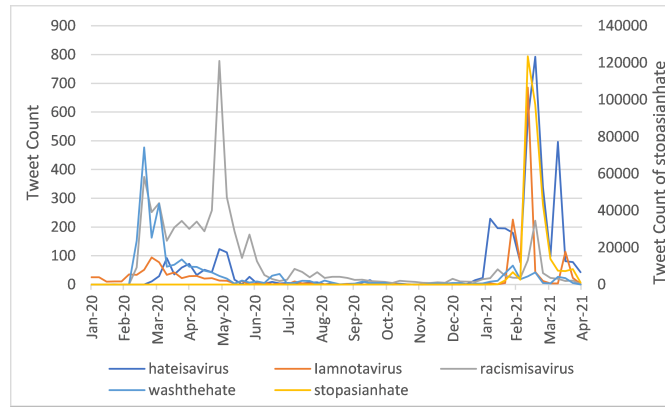


(b) Frequencies of Negative Keywords in India

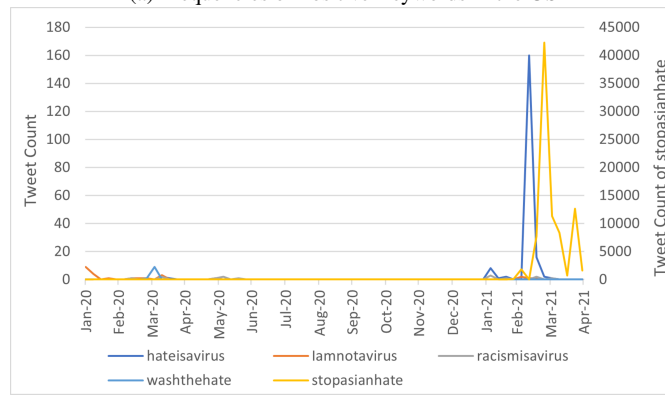


(c) Frequencies of Negative Keywords in Brazil

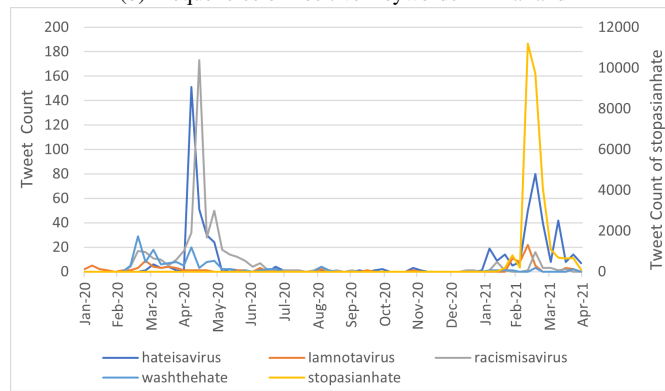
Fig. 14 Frequencies of each negative keyword over the data collection period created by the top three countries of production.



(a) Frequencies of Positive Keywords in the US

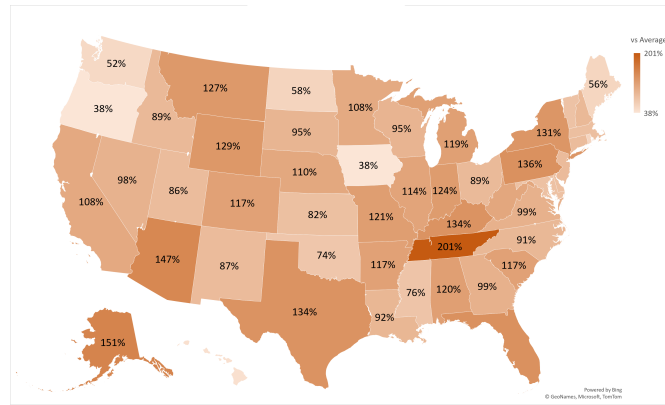


(b) Frequencies of Positive Keywords in Thailand

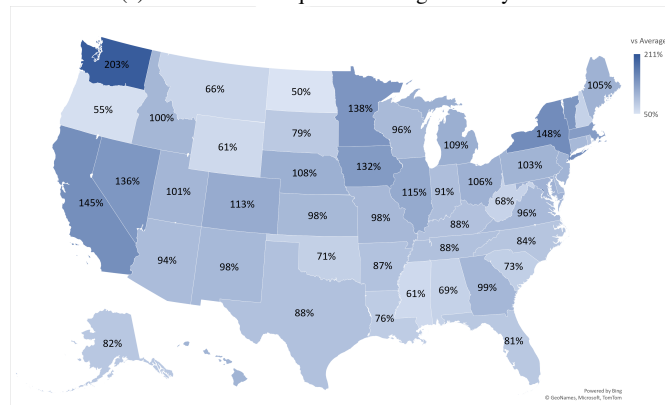


(c) Frequencies of Positive Keywords in Canada

Fig. 15 Frequencies of each positive keyword over the data collection period created by the top three countries of production.



(a) Normalized Frequencies - Negative Keywords



(b) Normalized Frequencies - Positive Keywords

Fig. 16 Normalized comparison of the frequencies of negative and positive keywords. The raw frequencies of the sample stream were first divided by the total number of tweets in the archive search produced in each state. Then, these ratios were averaged across all states.

containing positive keywords (CA = 45,886; NY = 48,595). As depicted in Fig. 15(a), in the U.S., the top producer of positive tweets, positive keyword use was low until February 2020, after which point it began to rise steadily. The most frequently used keyword was “stopasianhate,” followed by “racismisvirus” and “hateisavirus,” which reached their peaks in May 2020 and March 2021, respectively. “Washthehate” and “hateisavirus” showed surges in March 2020 and April 2021, respectively. Notably, the use of positive keywords was minimal from June to December 2020.

Fig. 15(b) shows that in Thailand, the second-highest producer of positive tweets, positive keyword use was almost non-existent until January 2021, but then picked up significantly from February 2021 onwards. The most commonly used keyword in Thailand was “stopasianhate,” which reached its peak in March 2021. Lastly, as depicted in Fig. 15(c), in Canada—the third top producer of positive tweets, the use of positive keywords was minimal until April 2020, but steadily increased thereafter.

“Stopasianhate” was once again the most frequently used term, as in the U.S. and Thailand. The frequency of this keyword peaked in March 2021 and then began to decline. Together, these findings demonstrate that different countries showed distinct patterns in the frequency of positive keywords over time. The widespread use of “stopasianhate” across all three countries highlights the urgency of addressing anti-Asian sentiment globally.

3.3 Analyses Using the 1% Dataset

The goal of this task was to normalize the frequencies of tweets based on the amount of overall Twitter activity in each state of the U.S. To this end, an initial ratio was computed by dividing the counts of positive and negative tweets with valid geolocation in the archive dataset by the count of tweets in the 1% dataset for each state. Then, the average initial ratio was computed across all states (i.e., 4.277 for positive keywords and 3.786 for negative keywords). A new ratio was calculated for each state by dividing the initial ratio by the average ratio. The final ratios for negative and positive keywords are reported in Fig. 16. As depicted, Tennessee had the highest ratio of negative keywords (i.e., 201% higher than the average), followed by Alaska, which was 151% higher than the average. In contrast, Washington D.C. had the highest ratio of positive keywords (i.e., 211% higher than the average), followed by Washington state (i.e., 203% higher than the average) and New York (i.e., 148% higher than the average).

3.4 Burst Analysis

We used Kleinberg’s burst analysis algorithm [52] to identify bursts of heightened negative and positive keyword use across time. This approach identifies bursts of activity in a series of events by modeling the transitions between two states—baseline and bursty. Bursty states are associated with periods of time when an event (e.g., negative or positive tweets) is unusually frequent. The approach uses two main parameters, s and $gamma$, which affect different aspects of the way the algorithm detects bursts.

- s : This parameter controls the threshold of event frequencies, or intensiveness, for each state. Higher values of this parameter will require stronger increases of activity to detect a burst.
- $gamma$: Gamma controls the difficulty of changing states. Higher values of this parameter will require more effort to switch states.

Multiple s and $gamma$ parameters for determining the sensitivity of the bursts were assessed in an iterative fashion. For example, as the s parameter cannot be less than or

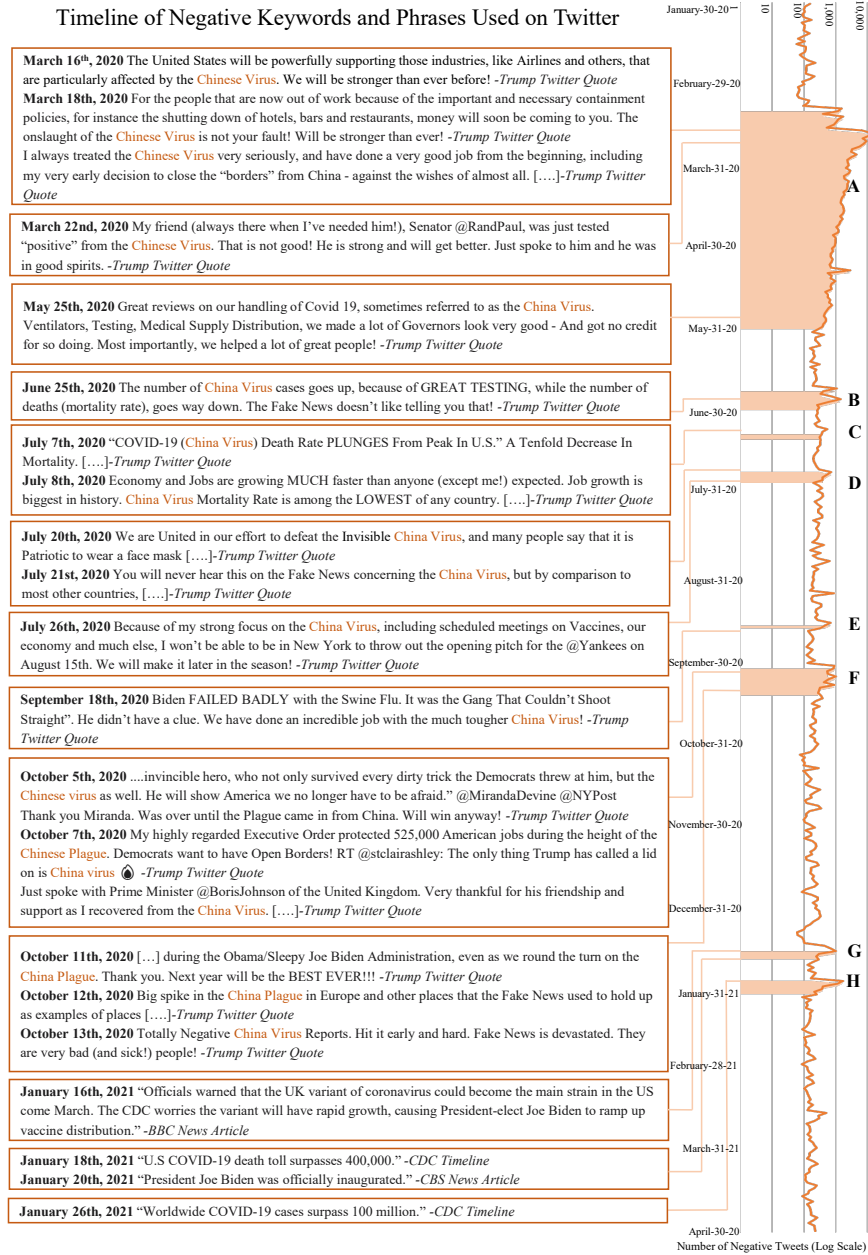


Fig. 17 Timeline of negative keywords and phrases with a subset of representative tweets occurring within bursts of heightened anti-Asian activity.

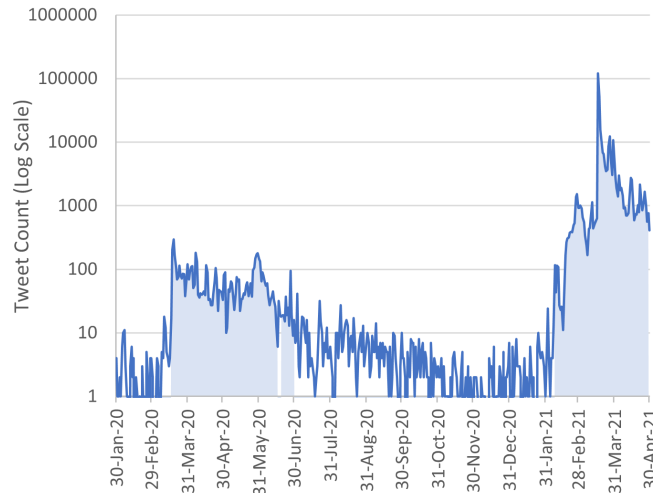


Fig. 18 Bursts of heightened positive activity.

equal to 1, steadily decreasing values of γ were tested ranging from the default of 1 to 0. During many of these combinations of s and γ values, either the analysis resulted in a binary burst (i.e., all of the data represents a burst of activity or none of the activity is considered a burst) or the bursts were inconsequential. From this testing, values of 1.1 for the s parameter and 0.0 for γ were selected. These values provided optimal visual output, in terms of producing the clearest separation between burst and non-burst regions and using multiple levels to capture burst intensity [52].

We performed separate burst analyses (with the same parameters) for the datasets with negative and positive keywords. Specifically, we used the burst detection algorithm to identify bursts in discrete bundles of events, where each bundle was defined as the set of negative or positive tweets received in a single day. For this analysis, we considered the tweets in the U.S. based on the geolocation labeling strategy previously described. To facilitate the processing of large frequency values using the Python Burst Detection library, we applied a logarithmic transformation before feeding the data to the algorithm. The output of this burst analysis step was a set of date ranges for the identified bursts.

Negative Keyword Use. The dates identified in the burst analysis were labeled with events on a timeline corresponding to the use of anti-Asian terminology (e.g., “Chinese virus,” “China virus”) on President Trump’s Twitter account, key political events, and COVID-19 milestones. In total, 8 bursts of activity were identified (labeled A through H in Fig. 17). Events were taken from dates up to 2 days before and after the beginning and end of the date ranges identified by the burst analysis. Events for Bursts A through F correspond primarily with tweets posted by (and originating from) Trump’s Twitter account [56]. Events for Bursts G and H were taken from

news media coverage of significant events [57, 58] that occurred at the time, as well as from the CDC's COVID-19 pandemic timeline [59].

Positive Keyword Use. Evaluating positive keyword use, 3 bursts of activity were identified, ranging from March 17, 2020 to June 16, 2020; June 19, 2020 to June 30, 2020; and February 2, 2021 to April 4, 2021 (Fig. 18). These bursts in positive keyword use immediately followed increases in physical violence and offline hate toward Asian Americans. For example, from March to June 2020, the Federal Bureau of Investigation reported increases in crimes directed toward Asian Americans [60]. Further, the burst of positive activity following February 2, 2021 culminates when there was a marked increase in physical violence against Asian individuals. For example, within this time frame, the highly-publicized Atlanta-area spa shootings occurred, in which Asian women were targeted, leading to the deaths of 8 individuals[51]. There were also several reports of individuals of Asian-descent being verbally and physically assaulted in public, resulting in serious injury or death [61, 62]. The burst in positive keyword use, in the form of prosocial, counter-hate messages, could be interpreted as a protective response to raise awareness as protests, rallies, and non-profit organizations were mobilized to fight this hostility [63, 64, 65]. Additionally, during this time (i.e., on March 23, 2020), President Trump publicly denounced his use of the term 'Chinese virus' and stated that Asian Americans should "not be blamed in any way, shape, or form" for the pandemic [66, 67].

4 Discussion and Concluding Remarks

The present study investigated temporal and geographic trends in anti-Asian prejudice and counter-hate messages on Twitter in the 15 months after the World Health Organization declared COVID-19 a public health emergency. Consistent with other recent research, our findings indicate that the increased prevalence of anti-Asian prejudice during the early stages of the pandemic was a global phenomenon [68, 69]. Specifically, marked increases in anti-Asian hate on Twitter occurred during the months of February 2020 and March 2020. Strong similarities in the trends for the overall data and for the data from the top 10% of producers provide some evidence that a relatively small but influential number of users were contributing largely to anti-Asian and counter-hate discussions online. Further, dis-aggregation of the data by type of tweet (e.g., original tweet vs. retweet) provided a more nuanced look at how and the extent to which anti-Asian and counter-hate messages are shared. The finding that replies were more likely to contain negative content might be an indication of a call to discuss anti-Asian hate online, given the interactive nature of replying. In contrast, retweets were more likely to contain positive content. Because retweeting may reflect the motive to share, promote, or express agreement with an original tweet, this finding may be indicative of greater support for positive content compared to negative content on Twitter [39]. Although we did not have access to information regarding users' motives for retweeting, our interpretation is consistent

with prior research finding that retweets are a good indicator of links between like-minded individuals [70] and that retweeting content is associated with increased trust in, interest in, and agreement with the original message [71].

Our analyses also revealed geographic differences in the frequency of negative (anti-Asian) and positive (counter-hate) content generated by Twitter users, both within the U.S. and on a global scale. Although we did not directly assess Hofstede's cultural dimensions [30], our findings are largely aligned with this theoretical framework. That is, Hofstede has indicated that the U.S. is high on individualism, India is high on power distance, and Brazil is high on uncertainty avoidance. This is particularly insightful, as it lends support to Ng's [31] findings that countries high on these factors might have a greater inclination to attribute blame to others in the face of uncertain events. The spread of positive keywords was more geographically focused in the U.S. Because individuals around the world increasingly view the U.S. as a source of influence [72], greater positivity (e.g., anti-racist messages) in online discourse within the U.S. may be beneficial for increasing positive online discourse on a global scale.

New York, California, and Florida were the largest producers of anti-Asian keywords within the U.S. in our archive dataset (based on our query of over 13 million tweets). However, a complementary analysis performed on a random sample of approximately 1% of publicly available tweets from a single date yielded additional insights. When considering the 1% of tweets on a given day, the states with the highest ratio of anti-Asian keywords to all Twitter content generated by users in the state were Tennessee and Alaska. In contrast, whereas California and New York were the largest producers of counter-hate keywords in our archive dataset, Washington D.C. had the highest ratio of counter-hate keywords in the 1% dataset, followed by Washington state and New York. The greater positive Twitter content generated by users in New York, in particular, is interesting in light of the relatively higher rate of crime targeting AAPI individuals in this state. That is, data from Stop AAPI Hate [73] indicates that out of 9,081 reported incidents of anti-Asian hate (i.e., physical violence, online harassment, civil rights violations) in the U.S. from March 2020 to June 2021, roughly 15% occurred in New York. The dynamic ways in which prejudice manifests itself in face-to-face interactions and online spaces—and the role of social media in conveying messages of support and solidarity in response to acts of racial animosity—warrant further empirical attention.

Using burst analysis, we identified several significant surges (i.e., bursts) in the frequency of both anti-Asian and counter-hate keywords on Twitter. Examination of these bursts in relation to relevant content generated by President Trump on Twitter, political events, and key milestones in the COVID-19 timeline helps contextualize these temporal findings and underscores the extent to which social media can both reflect and potentially influence anti-Asian sentiment. Crucially, our results are largely consistent with previous research indicating that President Trump's use of politically incorrect terminology when discussing political events precipitated increases in White nationalist ideals and racism [74], broadly, and the finding that bursts of negative activity occurred after President Trump started using anti-Asian rhetoric in his tweets, speeches, and interviews during the pandemic [9]. Furthermore, the

complexity of the prejudice fueled by and evident throughout the pandemic is perhaps illustrated by the political connotation of some of the anti-Asian keywords. For example, “ccpvirus”—in reference to the Chinese Communist Party—likely stemmed from news reports that this political party withheld information about COVID-19 during the early months of the pandemic [75]. A better understanding of the link between these bursts of activity and offline behavior may be particularly beneficial for social media platforms. For example, following a hate-speech triggering event (e.g., a violent attack or hate crime), social media companies might closely monitor hateful keywords and evaluate the use of mechanisms to amplify anti-hate messaging.

Finally, although there were more transitions from positive to negative (counter-hate to anti-Asian) than from negative to positive (anti-Asian to counter-hate) content within Twitter replies, bursts in the frequency of positive messages following violent attacks directed at Asian individuals highlight the protective nature of counter-hate messages online. This heightened positive activity can bring increased awareness of anti-Asian prejudice through “hashtag activism.” It remains unclear, however, whether surges in the use of counter-hate keywords (e.g., #hateisavirus, #stopasian-hate) led to a measurable reduction in verbal and physical attacks against AAPI individuals; notably, similar campaigns aimed at reducing violence have lost momentum over time [37].

4.1 Limitations and Future Directions

A few noteworthy limitations of this research stem from the data sample. For example, only tweets containing one of the English keywords or hashtags were collected and subsequently analyzed. Although over 1.4 billion individuals speak English globally [76] and about 53% of all tweets are in English [77], given that the COVID-19 pandemic and violence directed at individuals of Asian-descent occurred on a global scale (including in regions where English is not a commonly-spoken language), our exclusive reliance on tweets written in English limits the generalizability of our results. Future research using culturally-relevant anti-Asian and counter-hate keywords in other languages will contribute to a more complete understanding of the spread of anti-Asian hate online during this time.

Furthermore, we only utilized the top 20 countries with the most Twitter users (as of July 2020) along with their 5 most populous cities as locations within our geolocation strategy. Although we were able to classify roughly 10% of the collected data using this strategy, approximately 50% of our data did not contain a valid geolocation and approximately 40% did not contain any geolocation at all. This decreased the usable sample size for our analyses and leaves open questions about related phenomena in smaller countries and nations with limited access to Twitter. Although research focusing on Canadian [78] and U.K. [29] populations and findings from other global analyses (e.g., [31]) are largely consistent with our results, future research should evaluate these geographical trends across a wider range of countries.

Finally, it is important to note that the patterns of findings discussed in this research are correlational in nature. It is thus not possible to conclude that hateful messages directed toward Asian individuals caused or were caused by increases in offline violent attacks, salient events, statements by prominent figures, etc. It is also not possible to infer that rises in COVID-19 cases or deaths caused increases in anti-Asian or counter-hate messages. These relationships could also be impacted by a variety of confounding factors, such as individuals' political beliefs, access to and willingness to receive the COVID-19 vaccine, knowledge about the pandemic, and coverage of violent attacks against Asian individuals in the news media. Interestingly, however, we note that some research has supported the effectiveness of using Twitter activity as an indicator of the "wisdom of crowds." Specifically, Turiel et al. [79] found that the intensity of COVID-19-related tweets in different geographic regions during incipient stages of the pandemic predicted COVID-19 morbidity rates in these regions one month later.

This research aims to build initial awareness of patterns of hate and counter-hate on social media, with important questions remaining regarding the mechanisms underlying these observations. Our hope is that our efforts to expand on recent research in this area will contribute to a deeper understanding of how prejudice and hatred, as well as empathy and counter-hate, proliferate online during global crises.

Acknowledgements This work was supported by the National Science Foundation under awards #2036127 and #2227488. Additionally, the authors would like to thank Johnny Hudson (Arizona State University) for his contributions to an early version of the Twitter timeline.

References

1. World Health Organization, "Timeline: WHO's COVID-19 response", <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/interactive-timeline>, 2021.
2. Google Trends, "Coronavirus search trends", https://trends.google.com/trends/story/US_cu_4Rjdh3ABAABMHM_en, 2021.
3. World Health Organization, "World health organization best practices for the naming of new human infectious diseases", https://apps.who.int/iris/bitstream/handle/10665/163636/WHO_HSE_FOS_15.1_eng.pdf, 2015.
4. S. Darling-Hammond, E. K. Michaels, A. M. Allen, D. H. Chae, M. D. Thomas, T. T. Nguyen, M. M. Mujahid, and R. C. Johnson, "After "the China virus" went viral: Racially charged coronavirus coverage and trends in bias against Asian Americans," *Health Education and Behavior*, vol. 47, ed. 6, pp. 870-879, 10.1177/1090198120957949, 2020.
5. A. R. Gover, S. B. Harper, and L. Langton, "Anti-Asian hate crime during the COVID-19 pandemic: Exploring the reproduction of inequality," *American Journal of Criminal Justice*, pp. 1-21, 10.1007/s12103-020-09545-1, 2020.
6. T. Levenson, "Stop trying to make "wuhan virus" happen," <https://www.theatlantic.com/ideas/archive/2020/03/stop-trying-make-wuhan-virus-happen/607786/>, 2020.
7. World Health Organization, "A guide to preventing and addressing social stigma," <https://www.who.int/docs/default-source/coronaviruse/covid19-stigma-guide.pdf>, 2020.

8. A. Smith, "Photo of Trump remarks shows 'corona' crossed out and replaced with 'Chinese' virus", *NBC News*, 2020, <https://www.nbcnews.com/politics/donald-trump/photo-trump-remarks-shows-corona-crossed-out-replaced-chinese-virus-n1164111>
9. Y. Hswen, X. Xu, A. Hing, J. B. Hawkins, J. S. Brownstein, and G. C. Gee, "Association of #COVID19 versus #Chinesevirus with anti-Asian sentiments on Twitter: March 9–23, 2020," *American Journal of Public Health*, vol. 111, ed. 5, pp. 956-964, 2021.
10. N. G. Ruiz, K. Edwards, and M. H. Lopez, "One-third of Asian Americans fear threats, physical attacks and most say violence against them is rising," *Pew Research Center*, <https://www.pewresearch.org/fact-tank/2021/04/21/one-third-of-asian-americans-fear-threats-physical-attacks-and-most-say-violence-against-them-is-rising/>, 2021.
11. B. Levin, and A. Grisham, "Fact sheet: Anti-Asian prejudice March 2021," *Center for the Study of Hate and Extremism*, <https://www.csusb.edu/sites/default/files/FACT%20SHEET-%20Anti-Asian%20Hate%202020%20rev%203.21.21.pdf>, 2021.
12. M. Costello, L. Cheng, F. Luo, Ho. Hu, S. Liao, N. Vishwamitra, M. Li, and E. Okpala, "COVID-19: A pandemic of anti-Asian cyberhate", *Journal of Hate Studies*, vol. 17 ed. 1, pp. 108–118, 10.33972/jhs.198, 2020.
13. H. Tessler, M. Choi, and G. Kao, "The anxiety of being Asian American: Hate crimes and negative biases during the COVID-19 pandemic," *American Journal of Criminal Justice*, vol. 45, ed. 4, pp. 636-646, 10.1007/s12103-020-09541-5, 2020.
14. T. T. Nguyen, S. Criss, P. Dwivedi, D. Huang, J. Keralis, E. Hsu, L. Phan, L. H. Nguyen, I. Yardi, M. M. Glymour, A. M. Allen, D. H. Chae, G. C. Gee, and Q. C. Nguyen, "Exploring U.S. shifts in anti-Asian sentiment with the emergence of Covid" *International Journal of Environmental Research and Public Health*, vol. 17, ed. 19, 10.3390/ijerph171970323, 2022.
15. Stop AAPI Hate, "Stop AAPI hate national report 3.19.20 - 8.5.20," https://www.asianpacificpolicyandplanningcouncil.org/wp-content/uploads/STOP_AAPI_Hate_National_Report_3.19-8.5.2020.pdf 2020.
16. J.Cao, C. Lee, W. Sun, and J.C. De Gagne, "The #StopAsianHate movement on Twitter: A qualitative descriptive study", *International Journal of Environmental Research and Public Health*, vol. 19(7),3757, 10.3390/ijerph19073757, 2022.
17. Harrison, R. and Hamilton, K., "Hate speech on Twitter predicts frequency of real-life hate crimes", *NYU*, <https://www.nyu.edu/about/news-publications/news/2019/june/hate-speech-on-twitter-predicts-frequency-of-real-life-hate-crime.html>, 2019.
18. A. Toliyat, S. I. Levitan, Z. Peng, and R. Etemadpour, "Asian hate speech detection on Twitter during COVID-19", *Frontiers in Artificial Intelligence*, vol. 5, 10.3389/frai.2022.932381, 2022.
19. Anti-Defamation League, "ADL report: Anti-Asian hostility spikes on Twitter after President Trump's COVID diagnosis", <https://www.adl.org/news/press-releases/adl-report-anti-asian-hostility-spikes-on-twitter-after-president-trumps-covid-2020>.
20. E. Mcguire, "Anti-Asian hate continues to spread online amid COVID-19 pandemic", <https://www.aljazeera.com/news/2020/04/anti-asian-hate-continues-spread-online-covid-19-pandemic-200405063015286.html>, 2020.
21. A. D. Dubey, "The resurgence of cyber racism during the COVID-19 pandemic and its after effects: Analysis of sentiments and emotions in Tweets," *JMIR Public Health and Surveillance*, vol. 6, ed. 4, 10.2196/19833, 2020.
22. J. Y. Kim, and A. Kesari, "Misinformation and hate speech: The case of anti-Asian hate speech during the COVID-19 pandemic", *Journal of Online Trust and Safety*, vol. 1, ed. 1, 10.54501/jots.v1i1.13, 2021.
23. F. Miro-Llinares, and J.J. Rodrigez-Sala, "Cyber hate speech on twitter: Analyzing disruptive events from social media to build a violent communication and hate speech taxonomy", *WIT Press*, vol. 11, pp. 406 - 415, 10.2495/DNE-V11-N3-406-415, 2016.
24. B. Kim, E. Cooks, and S.-K. Kim, "Exploring incivility and moral foundations toward Asians in English-speaking tweets in hate crime-reporting cities during the COVID-19 pandemic", *Internet Research*, vol. 32, ed. 1, pp. 362-378, 10.1108/INTR-11-2020-0678, 2020.
25. J.An, H. Kwak, C. S. Lee, B. Jun, Y-Y. Ahan, "Predicting anti-Asian hateful users on Twitter during COVID-19", *arXiv*, vol. 1, 10.48550/arXiv.2109.07296, 2021.

26. H. Lyu, Y. Fan, Z. Xiong, M. Komisarchik and J. Luo, "Understanding public opinion toward the #StopAsianHate movement and the relation with racially motivated hate crimes in the US", *IEEE Transactions on Computational Social Systems*, vol. 10, no. 1, pp. 335-346, doi: 10.1109/TCSS.2021.3136858, 2023.
27. S. O. Cheng, "Xenophobia due to the coronavirus outbreak – A Letter to the editor in response to "the socio-economic implications of the coronavirus pandemic (COVID-19): A review"", *Int J Surg*, vol. 79, pp. 13-14, 10.1016/j.ijsu.2020.05.017, 2020.
28. M. Coates, "Covid-19 and the rise of racism." *BMJ*, vol. 06, 369:m1384, 2020.
29. S. Schumann, and Y. Moore, Y. "The COVID-19 outbreak as a trigger event for sinophobic hate crimes in the United Kingdom." *The British Journal of Criminology*, vol. 63, pp. 367-383, 10.1093/bjc/azac015, 2023.
30. Hofstede Insights, "Tool to compare countries", <https://www.hofstede-insights.com/hi/product/compare-countries/>, 2023.
31. R. Ng, "Anti-Asian sentiments during the COVID-19 pandemic across 20 countries: Analysis of a 12-billion-word news media database", *Journal of Medical Internet Research*, vol. 8, ed. 23, 10.2196/28305, 2021.
32. W. Sun, "Introduction to the Special Issue: "Media and communication during the COVID-19 (coronavirus) pandemic", *Howard Journal of Communications*, vol. 32, ed.5, pp. 415-420, 10.1080/10646175.2021.1966857, 2021.
33. L. Y. Dhanani and B. Franz, "Unexpected public health consequences of the COVID-19 Pandemic: A national survey of anti-Asian attitudes in the USA" *Int J Public Health*, vol. 65, ed. 6, pp. 747-754, doi: 10.1007/s00038-020-01440-0, 2020.
34. R. Goodwin, S. Haque, F. Neto, and L. B. Myers, "Initial psychological responses to influenza A, H1N1 ("swine flu")". *BMC Infectious Diseases*, vol. 9, ed. 166, doi: 10.1186/1471-2334-9-166, 2009.
35. M. S. Khan, A. Osei-Kofi, A. Omar, H. Kirkbride, A. Kessel, A. Abbara, D. Heymann, A. Zumla, and O. Dar, "Pathogens, prejudice, and politics: The role of the global health community in the European refugee crisis", *Lancet Infect Dis*, vol. 16, pp. e173–e177, doi: 10.1016/S1473-3099(16)30134-7, 2016.
36. M. McCauley, S. Minsky, and K. Viswanath, "The H1N1 pandemic: media frames, stigmatization and coping", *BMC Public Health*, vol. 13, pp. 1116, doi: 10.1186/1471-2458-13-1116, 2013.
37. B. Wheeler, S. Jung, M.C. Barioni, M. Purohit, D. Hall, and Y. Silva, "#WashTheHate: Understanding the prevalence of anti-Asian prejudice on Twitter during the COVID-19 pandemic", *Proceedings of the 2022 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 10.1109/ASONAM55673.2022.10068578, 2022.
38. S. Lindgren, "Movement mobilization in the age of hashtag activism: Examining the challenge of noise, hate, and disengagement in the #MeToo campaign", *Policy & Internet*, vol. 11, ed. 4, pp. 418-438, 10.1002/poi3.212, 2019.
39. D. Boyd, S. Golder and G. Lotan, "Tweet, tweet, retweet: Conversational aspects of retweeting on twitter", *43rd Hawaii International Conference on System Sciences*, Honolulu, HI, USA, pp. 1-10, doi: 10.1109/HICSS.2010.412, 2010.
40. K. Solovev, N. Prollochs, "Hate speech in the political discourse on social media: Disparities across parties, gender, and ethnicity", *Association for Computing Machinery*, pp. 3656–3661, 10.1145/3485447.3512261, 2022.
41. C. Rivers, B. Lewis, and S. Marmagas. "A framework for ethical use of Twitter for public health research", vol. 11, 2013.
42. Twitter. "Developer terms: Developer policy", <https://developer.twitter.com/en/developer-terms/policy>, 2022.
43. Twitter, "Developer Twitter API," *Twitter Developer*, <https://developer.twitter.com/en/docs/twitter-api>, 2021.
44. B. He, C. Ziems, S. Soni, N. Ramakrishnan, D. Yang, and S. Kumar, "Racism is a virus: Anti-Asian hate and counterspeech in social media during the COVID-19 crisis", *Proceedings of the 2021 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 10.1145/3487351.3488324, 2021.

45. A. Chiu, "Trump has no qualms about calling coronavirus the 'Chinese Virus.' that's a dangerous attitude, experts say", *Washington Post*, <https://www.washingtonpost.com/nation/2020/03/20/coronavirus-trump-chinese-virus/>, 2020.
46. E. Shim, "Asian Americans take campaign against 'Kung Flu' slur to the streets", *UPI*, https://www.upi.com/Top_News/US/2020/06/27/Asian-Americans-take-campaign-against-Kung-Flu-slur-to-the-streets/8321593296741/, 2020.
47. M. D. Cava, and K. Lam, "Coronavirus is spreading and so is anti-Chinese sentiment and xenophobia", *USA Today*, <https://www.usatoday.com/story/news/nation/2020/01/31/coronavirus-chinese-xenophobia-racism-misinformation/2860391001/>, 2020.
48. J. Jiang, E. Chen, S. Yan, K. Lerman, and E. Ferrara, "Political polarization drives online conversations about Covid-19 in the United States", *Human Behavior and Emerging Technologies*, vol. 2, ed. 3, pp. 200–211, 10.1002/hbe2, 2020.
49. Twitter, "Volume Streams," *Twitter Developer*, <https://developer.twitter.com/en/docs/twitter-api>, 2021.
50. A. Cohen. "Fuzzy string matching in Python", <https://pypi.org/project/fuzzywuzzy/>, 2022.
51. R. Fausset, N. Bogel-Burroughs, and M. Fazio. "8 dead in Atlanta spa shootings, with fears of anti-Asian bias", *New York Times*, <https://www.nytimes.com/live/2021/03/17/us/shooting-atlanta-acworth>, 2021.
52. J. Kleinberg, "Bursty and hierarchical structure in streams," *Data Mining and Knowledge Discovery*, vol. 7, ed. 4, pp. 373-397, 10.1023/A:1024940629314, 2003.
53. Statistica, "Number of Active Twitter Users in Selected Countries," *Statistica*, <https://developer.twitter.com/en/docs/twitter-api>, 2021.
54. World Population Review, "Largest Cities by Population," *World Population Review*, <https://worldpopulationreview.com/world-cities>, 2023.
55. World Health Organization, "WHO COVID-19 Dashboard," *World Health Organization*, <https://covid19.who.int/data>, 2023.
56. Trump Twitter Archive. "Trump twitter archive V2", <https://www.thetrumparchive.com>, 2021.
57. BBC US Canada. "Covid: UK variant could drive 'rapid growth' in US cases, CDC warns", <https://www.bbc.com/news/world-us-canada-55684878>, 2021.
58. S. Becket, G. Segers, K. Watson, M. Quinn, and C. Linton, "President Biden takes office, moving quickly to implement agenda", <https://www.cbsnews.com/live-updates/biden-inaugurated-46th-president-united-states/>, 2021.
59. Centers for Disease Control and Prevention. "CDC museum Covid-19 timeline" <https://www.cdc.gov/museum/timeline/covid19.html>, 2021.
60. Federal Bureau of Investigation, "Hate Crime," *Crime Data Explorer*, <https://crime-data-explorer.fr.cloud.gov/pages/explorer/crime/hate-crime>, 2023.
61. K. Tran, "2 elderly Asian women punched in the head in separate attacks on NYC subway", *Yahoo News*, https://news.yahoo.com/2-elderly-asian-women-punched-235848778.html?soc_src=social-sh&soc_trk=ma, 2021.
62. W. G. Kantor, "Filipino American man recounts brutal attack with box cutter on N.Y.C. subway: 'nobody helped'.", *People, Meredith Corporation*, <https://people.com/crime/filipino-american-man-recounts-brutal-attack-with-box-cutter-on-n-y-c-subway-nobody-helped/>, 2021.
63. A. V. Lozano, "People across U.S. protest anti-Asian hate following deadly spa shootings", *NBC News*, <https://www.nbcnews.com/news/us-news/people-across-u-s-protest-anti-asian-hate-following-deadly-n1261677>, 2021.
64. A. Rogin and A. Nawaz, "We have been through this before. Why anti-Asian hate crimes are rising amid coronavirus", *PBS News Hour*, <https://www.pbs.org/newshour/nation/we-have-been-through-this-before-why-anti-asian-hate-crimes-are-rising-amid-coronavirus>, 2020.
65. Stop AAPI Hate, "About" <https://stopaapihate.org/about/>, 2022.
66. K. Leigh, "Trump says he'll stop using the term 'Chinese Virus'", *Bloomberg*, <https://www.bloomberg.com/news/articles/2020-03-25/trump-says-he-ll-stop-using-chinese-virus-easing-blame-game>, 2020.

67. B. Samuels, "Americans after repeatedly using term 'Chinese virus'", *THE HILL*, <https://thehill.com/homenews/administration/489107-trump-expresses-support-for-asian-americans-after-repeatedly-using>, 2020.
68. X. Tan, R. Lee, and L. Ruppanner, "Profiling racial prejudice during COVID-19: Who exhibits anti-Asian sentiment in Australia and the United States?", *Australian Journal of Social Issues*, vol. 6, ed. 5, 10.1002/ajs4.176, 2021.
69. R. Tekumalla, Z. Baig, M. Pan, L. A. R. Hernandez1, M. Wang, and J. M. Banda, "Characterizing anti-Asian rhetoric during the COVID-19 pandemic: A sentiment analysis case study on Twitter", *Proceedings of the ICWSM Workshops*, 10.36190/2022.81, 2022.
70. B. Evkoski, A. Pelicon, I. Mozetič, N. Ljubešić, and P. Kralj Novak, "Retweet communities reveal the main sources of hate speech", *PLoS One*, vol. 17 ed. 2, e0265602, 2022.
71. P. T. Metaxas, E. Mustafaraj, K. Wong, L. Zeng, M. O'Keefe, and S. Finn, "What do retweets indicate? Results from user survey and meta-review of research", *In Proceedings of The International AAAI Conference on Web and Social Media*, vol. 9, no. 1, pp. 658-661, 2015.
72. J. Gramlich and K. Devlin, "More people around the world see U.S. power and influence as a 'major threat' to their country. *United States of America* <https://policycommons.net/artifacts/616940/more-people-around-the-world-see-us/1597687/>, 2019.
73. A. J. Y. Horse, R. Jeung, R. Lim, B. Tang, M. Im, L. Higashiyama, ... and M. Chen, "Stop AAPI hate national report", <https://stopaapihate.org/wp-content/uploads/2021/08/Stop-AAPI-Hate-Report-National-v2-210830.pdf>, 2020.
74. J. G. Shafer, "Donald Trump's political incorrectness: Neoliberalism as frontstage racism on social media", *Social Media + Society*, 10.1177/2056305117733226, 2017.
75. C. Buckley and S. L. Myers, "As new coronavirus spread, China's old habits delayed fight" *New York Times*, <https://www.nytimes.com/2020/02/01/world/asia/china-coronavirus.html>, 2020.
76. Statista, "The most spoken languages worldwide in 2023", <https://www.statista.com/statistics/266808/the-most-spoken-languages-worldwide/>, 2023.
77. SemioCast, "SemioCast – Top languages on Twitter-stats", <https://semioCast.com/top-languages-on-twitter-stats/>, 2023.
78. S. Chen, and C. Wu, "#StopAsianHate: Understanding the global rise of anti-Asian racism from a transcultural communication perspective", *Journal of Transcultural Communication*, vol.1, ed. 1, pp. 5-19. 10.1515/jtc-2021-2002, 2022.
79. J. Turiel, D. Fernandez-Reyes, and T. Aste, "Wisdom of crowds detects COVID-19 severity ahead of officially available data", *Scientific Reports*, vol. 11, 13678 10.1038/s41598-021-93042-w, 2021.