A STUDY ON THE EFFECTS OF U.S. EVENTS ON SENTIMENT RELATING TO

MINORITY GROUPS ON TWITTER

by

Caleb Lane Wright

Honors Thesis

Appalachian State University

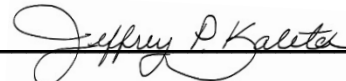Submitted to the Department of Computer Information Systems

And the honors college

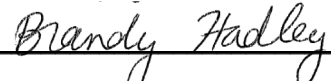In partial fulfillment of the requirements for the degree of

Bachelor of Science and Arts (BSBA)

May 2021

Approved by:
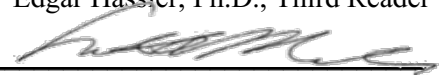
_____
Jeffrey Kaleta, Ph.D., Thesis Director

_____
Brandy Hadley, Ph.D., Second Reader

_____
Edgar Hassler, Ph.D., Third Reader

_____
Lorilee Medders, Ph.D., Departmental Honors Director

Jefford Vahlbusch, Ph.D., Dean, The Honors College

Abstract

This is a study on how U.S. national events affect the sentiment relating to different minority groups on Twitter. First, several groups of minorities were identified for examination. Next, the relationship between hashtag usage and Tweet composition relating to various minority groups was investigated. Using the Twitter API and several Python packages, we were able to gather Tweets relating to our areas of interest over the months of June to August 2020. Average Tweet sentiment was determined to see how that sentiment changed before, during, and after selected events. Three metrics were used to draw conclusions from the collected data. These were average sentiment of the Tweets, percentage of Tweets that were "retweets" versus those composed originally, and number of Tweets that were deemed "relevant" by twitter and therefore available for sample. We found that Tweets relating to the Black Lives Matter and related movements decreased in original tweet composition and average sentiment. Overall, this indicated a negative effect on the sentiment relating to Black Lives Matter initiatives during several related events. Tweets relating to the Coronavirus and people of Asian ethnicity did not change in average sentiment and decreased in original tweet composition in relation to the Coronavirus epidemic. Overall, this also indicated a negative effect on the sentiment relating to Asians due to this world event. Finally, Tweets relating to members of the LGBTQ+ community increased average sentiment, and decreased in original tweet composition during the month of June, commonly associated with "Pride Month". This indicated a positive effect on the sentiment relating to the LGBTQ+ community during this event. We can come to the overarching conclusion that U.S. national events do not have a direct effect on positive or negative sentiment independently, rather that these events cause a change in the sentiment that can be characterized as either more positive or more negative.

People are much more likely to retweet Tweets relating to events that elicit visceral reactions from people. This is a common sign of support. Since the original Tweet often has polarized sentiment due to the emotionally charged content, we can conclude that the events of the study will result in high retweet percentages with polarized sentiment.

Introduction

Social media's goal, at its core is to connect people over the internet. Before its invention, to find people with commonalities, people had to attend club meetings, subscribe to newsletters, and/or participate in very intentional and not always easy to find social gatherings. Since social media's introduction in society, it is now easier than ever to connect with people like you in even the most niche ways. Minority groups are no exception. Platforms like Facebook and Twitter allow minority groups to connect, share their experiences, raise support for causes important to them, and otherwise interact at a higher frequency than before in human history (Theocharis, Lowe, van Deth, and García-Albacete, 2015). These platforms also allow easy access for people who are not in minority groups to be exposed to all the aforementioned topics. Twitter has 330 million monthly active users largely based in the United States, but with activity from nearly every country in the world (Dean, 2021). Because of its vast scope and emphasis on using words to communicate, Twitter is a great place to get up-to-date information about how a large group of people feel about nearly any topic. Increased connection of minorities, paired with increased exposure of non-minorities to their experiences has created an environment where issues relating to minorities tend to have much more awareness, but are also more widely debated and criticized. A context of heightened awareness and debate on social media platforms means that what people are saying about these issues, including their biases, exaggerations, misconceptions, empathy, etcetera, has a real and weighty effect on how these issues are perceived and ultimately dealt with.

How people feel about issues, or their sentiment with regards to issues, has major implications on conversations surrounding these issues (Theocharis, Lowe, van Deth, and

García-Albacete 2015). As discussed, platforms like Twitter are a frictionless way for people to express their sentiment about topics affecting groups they may not even be a part of. This is good because it means people in power who are not part of minority groups can be publicly sympathetic or critical of an issue and ultimately raise awareness or be the catalyst of change (Theocharis, Lowe, van Deth, and García-Albacete, 2015). This also means that the general public can express their ideas about issues affecting minorities. In this regard, public support or criticism can greatly affect people's perceptions regarding these topics. If the general public's sentiment is overall supportive of an issue, then that issue is more likely to be looked at positively by people in power and more likely to enact a grassroots movement that helps create a change. If the general public's overall sentiment is negative with regards to an issue, then that issue is more likely to illicit condemnation from political figures and help shift public perception outside of social media negatively. In extreme cases, sentiment on Twitter can even be the catalyst of legislation affecting the issue at hand.

People's sentiment on an issue as expressed on Twitter can greatly affect the perception of an issue in a context much wider than a social media platform. This leads to the question; how do events related to minorities affect people's sentiment towards related conversation in social media? We know that social media sentiment can elicit change in the real world, but how does social media sentiment change when major world events happen? To help answer this question we will look at average sentiment, and the amount of uniquely created tweets before and after events to analyze how sentiment changes in accordance with major related events.

The remainder of the paper is organized as follows. First, we introduce relevant literature relating to social media and the use of data collected from Twitter. Also, we review

some of the methods available to analyze messages from social media sources. We then introduce our data collection methodology, analysis of the data, and a discussion of the results.

Literature Review

Social Media

Social media is a major form by which many of us communicate; how groups of people are perceived in social media often has direct effects on how they are perceived in the real world. Social media has a vast influence on our modern lives. If there is a better understanding of how people are seen on social media, injustices, misconceptions, and other problematic aspects of common perceptions can be opened up to conversation and addressed. To do this in a quantitative manner, this research examines which minority groups have the most negative and positive average sentiment on Twitter, and why.

Twitter is the social media platform with a heavy emphasis on words rather than pictures, like Instagram or Snapchat. Although pictures can tell a story about how a population perceives a people group, it is often done implicitly, and therefore is difficult to quantify (Lee, Lee, Moon and Sung, 2015). It is much easier to know what someone thinks about someone else when they can tell you using words.

Twitter has 330 million monthly active users largely based in the United States, but with activity from nearly every country in the world (Lin, 2020). Because of its vast scope and emphasis on using words to communicate, Twitter is a great place to get up-to-date information about how a large group of people feel about nearly any topic.

Social media, specifically Twitter, is a platform by which large amounts of information are diffused every day. Several studies have been done on the factors that affect information

diffusion on social media, however Stieglitz & Dang Xuan (2013) looked specifically at information diffusion as it relates to tweet sentiment. The context of this question was centered around political communication on Twitter. They found that Tweets with either high positive or negative sentiment were retweeted more often and more quickly than those that were more neutral in sentiment.

As shown by Stieglitz & Dang Xuan (2013), retweeting is a common response to Tweets that are emotionally charged. Retweeting is an easy way for Twitter users to disseminate information to their followers, but Suh, Hong, Pirolli and Chi (2010) wanted to find why some information spreads more widely than others. Based on a dataset of 74 million tweets, they identified common factors that were statistically significantly correlated with the rate at which tweets were retweeted and from this information, built a predictive retweet model. Two categories of features were analyzed: content features and context features. The content features that had a strong correlation with retweetability were URLs and hashtags. The context features that had a strong correlation with retweetability were the number of followers an account has and the age of the account. The past number of tweets was not significantly correlated with retweet rates. Stieglitz & Dang Xuan (2013) also looked at the effect of sentiment on the retweetabilty of tweets. Based on a data set of 64,431 political tweets, they found a positive correlation between more positive or negative sentiment in a tweet and how often that tweet was retweeted.

Although tweets can be retweeted to show support, Twitter is also used for organizing more tangible shows of support like protests and other forms of political action. Posting on Twitter is an avenue by which people sympathetic to a cause can communicate, mobilize and organize frictionlessly. Theocharis, Lowe, van Deth, and García-Albacete (2015) studied the

8

use of Twitter to do just that on protest events in Spain, Greece and the United States. They found that Twitter was frequently used for political discussion and to communicate information about these protests, while calls to action were in the minority.

As shown by Stieglitz & Dang Xuan (2013), hashtags increase the retweetability of tweets, but hashtags are also a way to find other tweets which relate to the hashtag's topic. Different racial and ethnic groups often use hashtags to relate to other members of their group. Sharma (2013) explored the effect of race-specific hashtags on "producing networked subjects which have the capacity to multiply the possibilities of being raced online." He finds that these hashtags play a major role in creating and preserving online identities for the communities that participate in race-specific hashtags.

Analyzing Social Media Messages

In language, each word has a sentiment attached to it based on context. Thanks to advances in computer science, there are tools like Vader Sentiment which have attached quantified sentiment analysis to words, statements, and even emojis based on context. Using these tools, it is possible to input a statement and get output with data about how positive, negative or neutral the statement is.

There are many available tools to conduct Sentiment analysis. Social media text is different from most other forms of online text because of the extremely high volume of content created as well as the tendency to use abbreviated language conventions to express sentiments. Because of the implicit nature of social media, a method attuned to the specific challenges related to analyzing the sentiment is necessary. Hutto and Gilbert (2014) created Vader, a Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text and compared how Vader performed against eleven typical state-of-practice benchmarks to test the

effectiveness of their sentiment analysis tool in the context of microblogging. They found that Vader outperforms individual human raters (F1 Classification Accuracy = .96 and .84, respectively). Hutto and Gilbert (2013) include the fact that often a comprehensive lexicon is required in order to accurately and quickly perform sentiment analysis. Linguistic Inquiry and Word Count (LIWC, pronounced "Luke"), is used in VADER and because it has been extensively validated, is easily expandable and inspectable.

Suh, Hong, Pirolli and Chi (2010) and Stieglitz & Dang Xuan (2013) have shown the use of hashtags and the more polarized the sentiment of the tweet, the more likely a tweet is to be retweeted and that this retweet is a sign of support. Paired with the idea sourced from Sharma (2013) that different ethnic groups use hashtags specific to their group, we can conclude that hashtags are a proficient way to find tweets relevant to different ethnic and racial groups. As shown by Theocharis, Lowe, van Deth, and García-Albacete (2015), tweets are also used to communicate about protests and other politically charged movements. Synthesizing these two conclusions, we can assume that tweets with ethnic-specific hashtags can be used to communicate about protests and other world events relating to that ethnic group. Using Hutto and Gilbert's (2014) VADER sentiment, we can gage the sentiment of these relevant tweets and analyze changes as world events relating to various minorities occur.

Research Questions

As aforementioned, research by Theocharis, Lowe, van Deth, and García-Albacete (2015) show that posting on Twitter is an effective way to organize tangible political movement. We want to see how the sentiment relates to these movements that often gain traction via Twitter as the movements progress. Specifically, we aim to answer the question

how do events related to minorities affect people's sentiment toward related conversations in social media?

During the height of the Black Lives Matter movement, the COVID-19 virus event and Pride month in 2020, we wanted to know exactly how these large-scale events affected the sentiment surrounding the minority groups most intimately involved with the aforementioned events. Our three key metrics, how large the relevant Tweet pool is, the average Tweet sentiment and the percentage of user-created Tweets, will give us an idea how sentiment on Twitter regarding these minority groups changes before, during and after the associated events.

Method

When starting this study our first logical step was to create a list of search terms representing minority groups of interest. The original aim was to take a holistic view of sentiment on Twitter regarding all minorities, thus a list of search terms was compiled which was later refined to the analysis of only a few representative groups. At the beginning of the study, we could not have foreseen national events relating to the Black Lives Matter movement or COVID-19 epidemic, so the direction of the analysis was adjusted to make use of these events as part of the study since they occurred during the time of the data collection process.

Once the minority groups were identified, we concluded that selecting appropriate Hashtags (on Twitter) was the best way to easily and accurately identify Tweets and search terms that related to those groups. Often when talking about a subject, users on Twitter will include a hashtag with their Tweet so that it can be easily identified and found by users who are interested in related topics. Because of the direct correlation between hashtags and the

11

topics discussed or opinions expressed within Tweet bodies, hashtags are a great way to identify Tweets relating to our selected minority groups. This is consistent with findings by Suh, Hong, Pirolli and Chi (2010), Stieglitz & Dang Xuan (2013), and Sharma (2013) that found that not only are retweets a sign of support of whatever the body of the Tweet is, but that ethnic and racial groups often have common hashtags that are used to identify other members in the group and find common experiences.

After establishing our method to accurately find Tweets talking about the minority groups of interest via hashtags, we needed to find hashtags specific to these groups. This was done via the website Hashtagsforlikes.co, which categorizes the most popular hashtags included in posts with a hashtag that you give it. As a starting point, the hashtag of the minority name was entered into the database and the most popular results were collected to ensure that we had as complete of a list as possible with all the hashtags that relate to a given minority group. As the study went on, our scripts had the flexibility to add other hashtags. Hashtags are very dynamic and reflect what is going on in the world, and therefore they change and there are new ones created constantly to identify Tweets of new issues and events. Because of the nature of hashtags, whenever a new event or issue began related to the minority groups of interest, we started collecting Tweets with hashtags that related to that event or issue.

Our next step was to establish a method to analyze the sentiment of our relevant Tweets. We chose to use a Python based sentiment analysis tool, Vader Sentiment Analyzer (Hutto and Gilbert, 2014). Vader Sentiment was concluded to be an effective way to analyze social media posts (Hutto and Gilbert, 2014) and is easy to insert into our Python scripts for data collection.

To collect data, we constructed a custom Python application, to search for and collect Tweets relating to our identified hashtags of interest. Data collection and compilation were handled with several Python packages (e.g. Tweepy, Pandas) and using the Twitter REST API. The Twitter REST API allows access to Tweets that are deemed the most relevant by Twitter's algorithms within the constraints provided. The sample of Tweets that are available are only those which are tweeted after you start your search query. One of the limitations of the Twitter's REST API is the window of time when tweets are collected. Twitter will only return 100 of the most relevant and most recent tweets over the past 7 days from a search. If there are more than 100 tweets in a given hour, then prior tweets are not collected. To overcome this, data collection requests made to Twitter occurred once per hour through the day. Because of this limitation, our available data for analysis was collected only over the time of our study.

To communicate directly with Twitter's REST API, a custom Python script was created using the Python library 'Tweepy'. Tweepy is a library that allows users to easily access the Twitter REST API and gather data. Using this package in the Anaconda web-based IDE, we developed a program that collects and archives relevant Tweets. When the Tweets are collected, they are processed in JavaScript Object Notation or JSON format. This format then needs to be adapted into a structured format that can be easily tabulated and translated into graphs and other representations. For this purpose, JSON data was stored in separate files in relation to the hashtag term that was searched. To this point, each JSON file was assigned a numeric code that corresponded with the hashtag contained in (the search term) the Tweet. For example, Tweets that related to Black minorities were identified with ID codes 10-100, 110-160 were assigned to Tweets related to Latinx hashtags, and Tweets related to Asians were assigned to IDs 170-220. Female gender groups were assigned 230-270 and LGBTQ+ were

assigned to 280-370. Hashtags related to minority groups with disabilities were assigned 380-410 and immigration groups were assigned 420-450. This coding system allowed us to group the tweets by corresponding minority group and easily identify which hashtag was included in the Tweet.

As each Tweet is put into a structured format, the sentiment is analyzed via Vader Sentiment (Hutto, C. J., & Gilbert, E., 2014). This allows us to extrapolate how positive, neutral, or negative the Tweet's sentiment is. Each Tweet is given four values. First, the Tweet is given three values (adding up to 1) indicating the percentage of neutral, positive, and negative words in a tweet. The Tweet is also given a composite sentiment value between negative one and one that indicates the sentiment of the overall tweet. According to Hutto and Gilbert (2014), this is a better measure of the Tweet's sentiment since this measure can take into consideration contextual clues like netspeak (e.g. LOL, ROTFL, Sux) and emoticons (e.g. 😊, 🙁) more effectively and adjust the score accordingly.

Finally, to assess the collected data, all data was converted to a single data file, which was analyzed using Tableau to organize and visualize the data. Details of the analysis are discussed in the following section.

Analysis and Results

Over the course of our study, the two types of data that were of interest to us were Tweet sentiment composition and the percentage of original tweets. For each of the three events studied, we gathered sentiment and retweet data through Tweets that were identified as relevant due to their hashtags. Whenever a new event occurs or there are significant developments in a current event, new hashtags are created to reflect those events or

14

developments. Because of this dynamic nature of hashtags, throughout the study, new hashtags relevant to our study were created and widely used. As we became aware of these hashtags, we started collecting data on them. Another note regarding our data collection: Over the course of data collection, we experienced an interruption from May 22nd to June 2nd, due to a power outage. However, the current analysis still provides us with ample data to see the effects of the events of interest.

Black Lives Matter Movement

Figure 1 shows the Tweet sentiment composition of the Black Lives Matter movement. We selected May 11th to June 17th as our time period of study for the movement because May 11th is 2 weeks before George Floyd was killed. Although the Black Lives Matter movement was prevalent before this date, because there was a significant increase in media coverage and participation from the general population following the death of George Floyd, we concluded this would be an appropriate start date for the study. This graph shows the average sentiment of Tweets relating to the Black Lives Matter movement during our given time period. The sentiment is presented along a scale from -1 to 1, where -1 reflects perfectly negative sentiment and 1 reflected perfectly positive sentiment. The #BlackLivesMatter and #BlackPower hashtags were prevalent before the start of the event, and therefore are our baselines to gauge the change (or lack thereof) in sentiment composition over the course of the event.

Black Tweet Comp
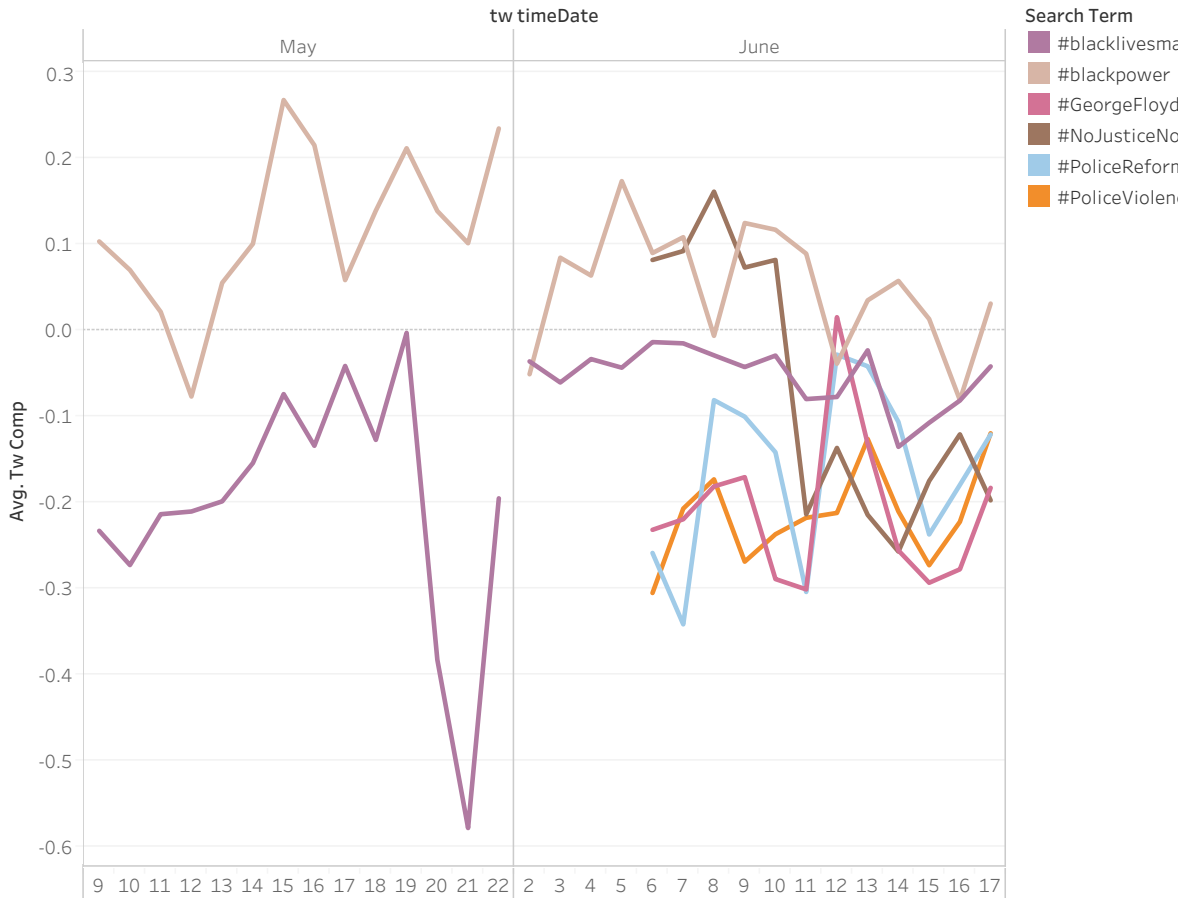
The trend of average of Tw Comp for tw timeDate Day broken down by tw timeDate Month. Color shows details about Search Term. The data is filtered on tw timeDate, which ranges from 5/9/2020 12:00:00 AM to 6/17/2020 11:59:59 PM and keeps Null values. The view is filtered on Search Term, which keeps 6 members.

Figure 1. Black Lives Matter Movement Tweet Sentiment Composition

As shown in Figure 1, almost all of the hashtags observed had average overall negative sentiment during some point in our selected time frame. Because the Black Lives Matter movement started due to an instance of violence, and there were weeks of protests and the general public was very angry with the government, it makes sense that the average sentiment of these Tweets would be negative. As shown by Theocharis, Lowe, van Deth, and García-Albacete (2015), political movements are often organized on Twitter. We can conclude that many of the Tweets relevant to the Black Lives Matter movement over the course of the event were likely used to organize protests, an emotionally charged event where people were

16

expressing their dissatisfaction for the United States Government's handling of police brutality. Because of the nature of these protests and the new use of the hashtags related to the movement, we can conclude that average sentiment of Tweets relevant to the Black Lives Matter movement decreases over the course of the two weeks after the event.

Figure 2 shows the percentage of original Tweets as opposed to retweets. If someone retweets an original Tweet, it counts as their Tweet but has a special flag to denote that the Tweet was not authored by the person who retweeted it. In Figure 2, we decided to use a different time constraint for the Black Lives Matter event due to data visualization constraints. From the figure we can conclude that there were significantly more retweets relevant to the Black Lives Matter Movement after the protests started.
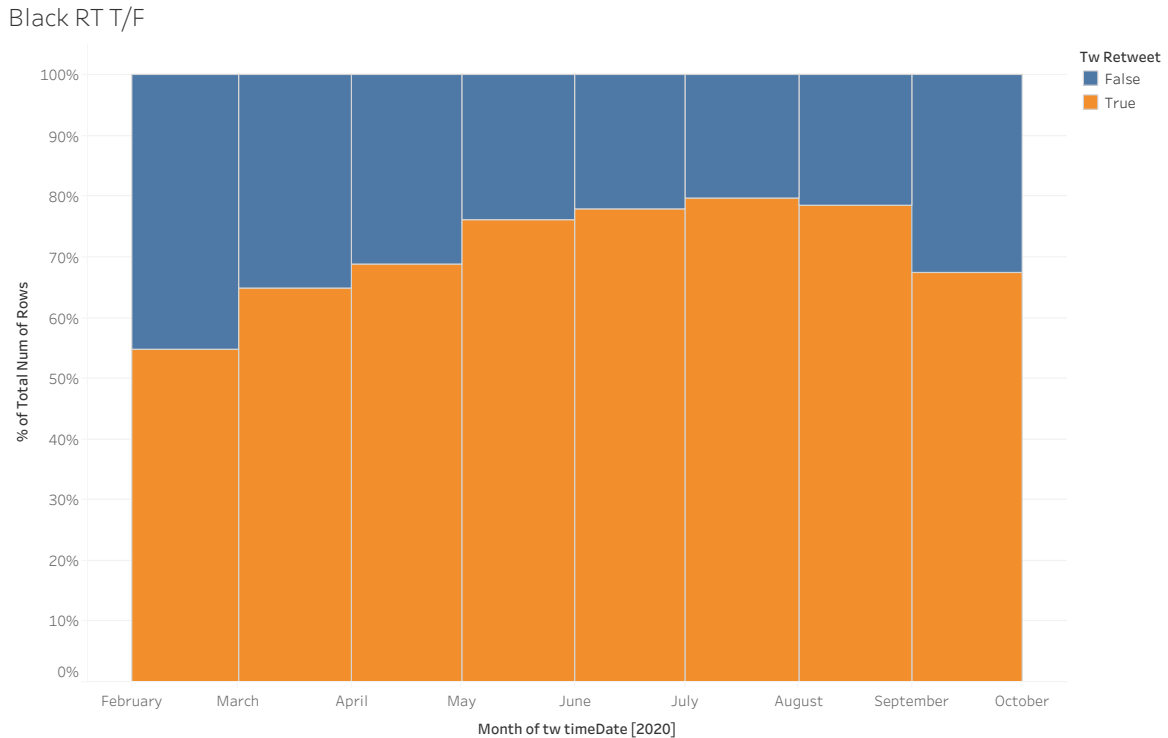
The plot of % of Total Num of Rows for tw timeDate Month. Color shows details about Tw Retweet. The data is filtered on tw timeDate Week and Search Term. The tw timeDate Week filter has multiple members selected. The Search Term filter keeps 6 members. Percents are based on each column of each pane of the table.

Figure 2. Black Lives Matter Movement Percentage of Original Tweets

Figure 2 shows that retweets were more prevalent after the start of the Black Lives Matter protests. This data supports the Stieglitz & Dang Xuan (2013) conclusion that retweeting is a common response to Tweets that are emotionally charged. The Black Lives Matter movement was an incredibly emotionally charged event and therefore people often reacted to their favorite celebrities or other notable people's Tweets related to this movement by retweeting their Tweets.

In conclusion, we found that Tweets relating to Black people during and after the Black Lives Matter protests decreased in original tweet composition and average sentiment. Overall, this indicated a negative effect on the sentiment relating to Black people during this event.

COVID-19

Figure 3 shows the average Tweet sentiment composition of the COVID-19 event over our selected time period. We selected the time frame of April 1st- May 15th because April was the first full month of data collection for the COVID-19 event. The time period in Figure 4 is the same as seen in Figure 3 due to the same logic. We can conclude from Figure 3 that average sentiment did not significantly change during the COVID-19 event. Sentiment of the relevant hashtags averaged slightly negative. A notable outlier is #chinaflu which had several positive spikes over the course of the study. These spikes can be explained by a lack of Tweets with this hashtag. Because the dataset isn't very big, as seen in figure 3, spikes are expected as each instance of a Tweet with the hashtag #chinaflu is going to have a more substantial influence on the average sentiment composition.
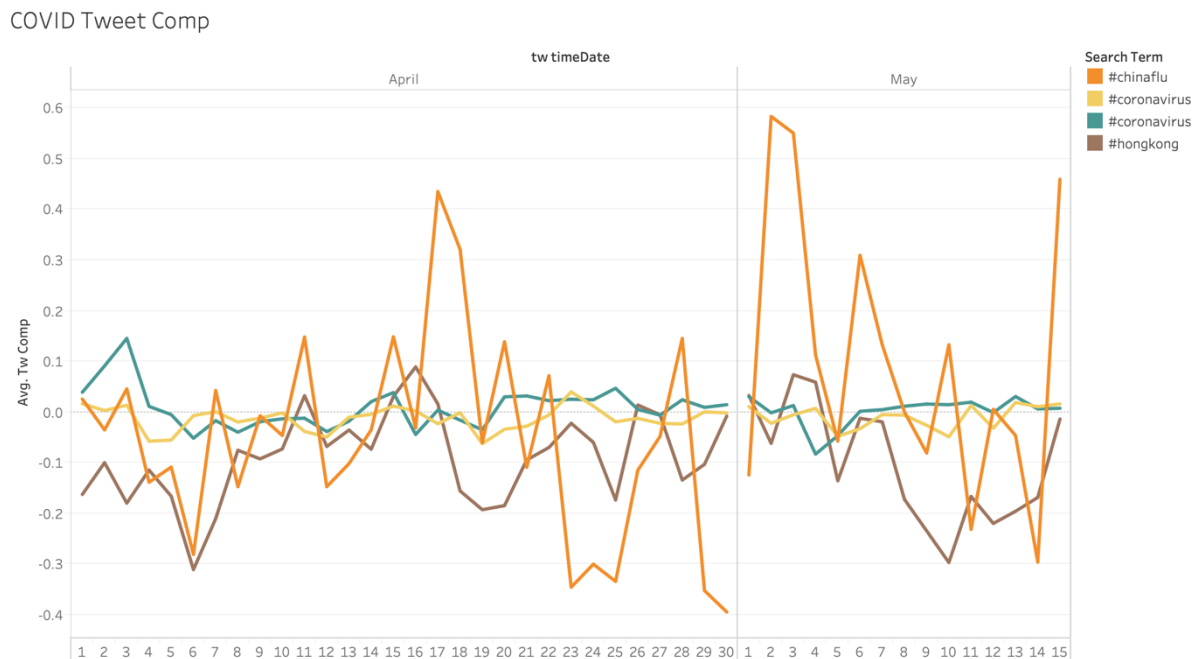
COVID Tweet Comp

The trend of average of Tw Comp for tw timeDate Day broken down by tw timeDate Month. Color shows details about Search Term. The data is filtered on tw timeDate, which ranges from 4/1/2020 12:00:00 AM to 5/15/2020 11:59:59 PM. The view is filtered on Search Term, which keeps #chinaflu, #coronavirus, #coronavirusoutbreak and #hongkong.

Figure 3. COVID-19 Tweet Sentiment Composition

Figure 3 shows a slightly negative average sentiment over the course of the COVID-19 event. This is expected of an event that is newsworthy and therefore widely objectively reported on meaning that there is little sentiment shown in Tweets. Tweets that were less informational, often Tweeted by individuals rather than news sources, can be expected to have negative sentiment on average because of the extensive negative effects of COVID-19 on the United States.

Figure 4 shows the percentage of original Tweets for the COVID-19 event over our selected time period. Once again, due to data visualization constraints, we chose a different event period of February 1st – October 31st, the full length of our study. As shown in figure 4, there was a significant increase in the number of retweets over the course of the study. In the

month of February, the percentage of relevant Tweets was 64.02%. During the course of the event there was an average of 72.86% of relevant Tweets that were retweets.
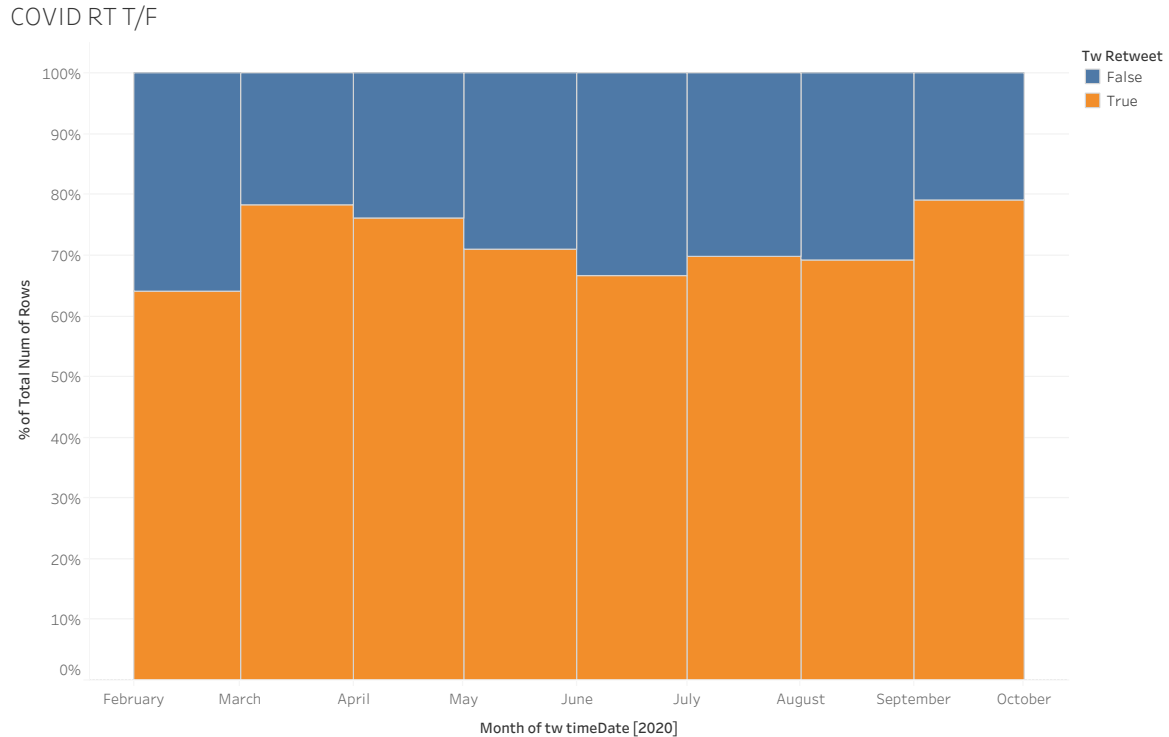
COVID RT T/F

The plot of % of Total Num of Rows for tw timeDate Month. Color shows details about Tw Retweet. The data is filtered on tw timeDate Week and Search Term. The tw timeDate Week filter has multiple members selected. The Search Term filter keeps #chinaflu, #chinaoutbreak, #coronavirus, #coronavirusoutbreak and #gay. Percents are based on each column of each pane of the table.

Figure 4. COVID-19 Percentage of Original Tweets

From Figure 4 we can conclude there were significantly more retweets during the COVID-19 event as opposed to before it. Since national lockdowns started in the United States in March, this data supports the Stieglitz & Dang Xuan (2013) conclusion that retweeting is a common response to Tweets that are emotionally charged. National lockdowns were an emotionally charged event and therefore people were likely to retweet information relating to the COVID-19 event, including Asian minorities to spread information and express agreement with opinions that, as shown in Figure 4, had an average negative sentiment.

In conclusion, there was no change in the average negative sentiment and a decrease in original tweet composition during the height of the Coronavirus epidemic (which still continues). The negative average sentiment supports our conclusion that polarized sentiment is the result of emotionally driven Twitter content. The decrease in original Tweet composition is also consistent with an event that draws visceral reactions from people. Because people want to show support for those who are experiencing the effects of the emotionally charged event, they are much more likely to retweet someone else's tweet than compose their own. Since emotionally charged events often have polarized sentiment, we see an increase in this polarization when there is an increase in retweets.

Pride Month

Figure 5 shows the Tweet sentiment composition during our selected time period for Pride Month. Pride Month is celebrated in June in the United States. Because of this, our selected time frame was one month before June and as far into June as we had data. From Figure 5, we can conclude there was a significant increase in the average sentiment composition related to the Pride Month event and members of the LGBTQ+ community. The most notable datapoints are the hashtags #pride and #loveislove. These two showed the greatest increase between the months of May and June.
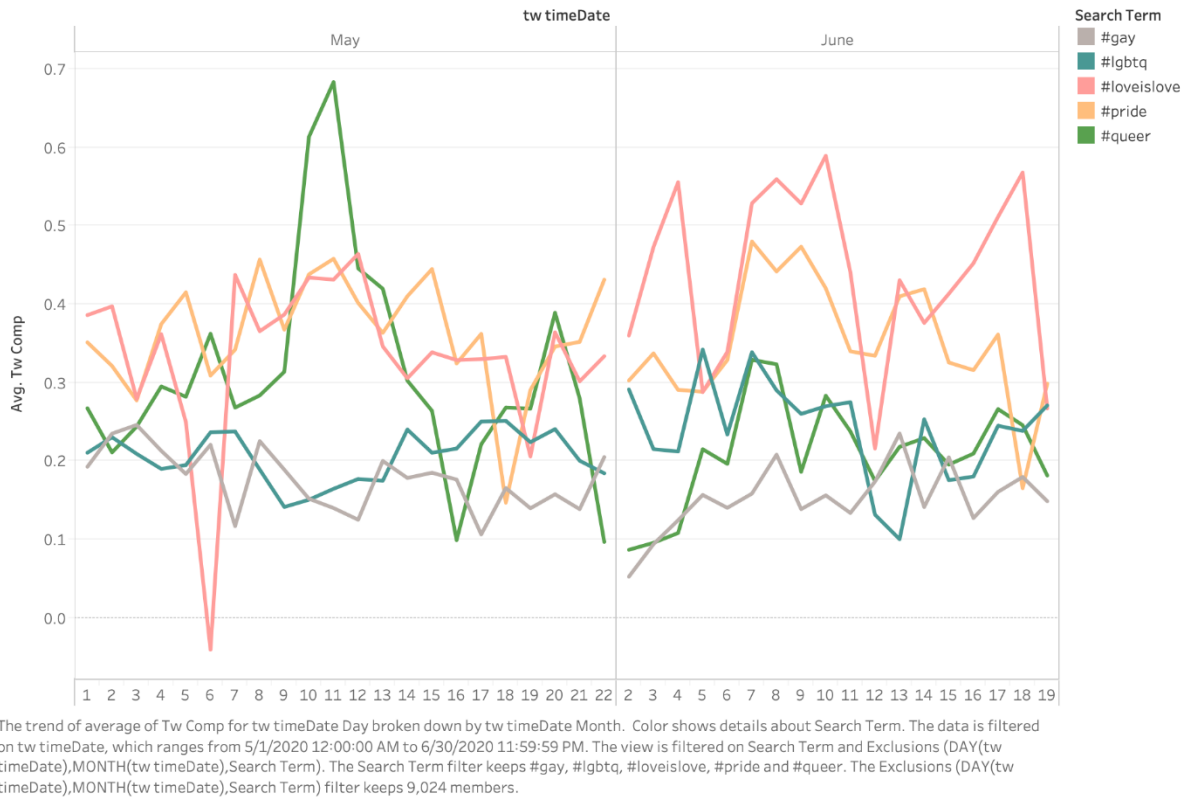
Figure 5. Pride Month Tweet Sentiment Composition

As the Pride Month event began, people were more likely to use Tweets with hashtags that related to the LGBTQ+ community and the average sentiment of these Tweets increased. Due to the nature of Pride Month, the idea that it is a celebration of the members of the LGBTQ+ community, it makes logical sense that the average sentiment related to Tweets about those in the community would increase during the event.

Figure 6 shows the percentage of relevant Tweets during the Pride Month event. It can be concluded from Figure 6 that the percentage of original Tweets decreased during the Pride Month event. During the month of May 59.49% of Tweets related to the LGBTQ+ community were retweets. During the month of June the number of retweets increased to 67.53%.
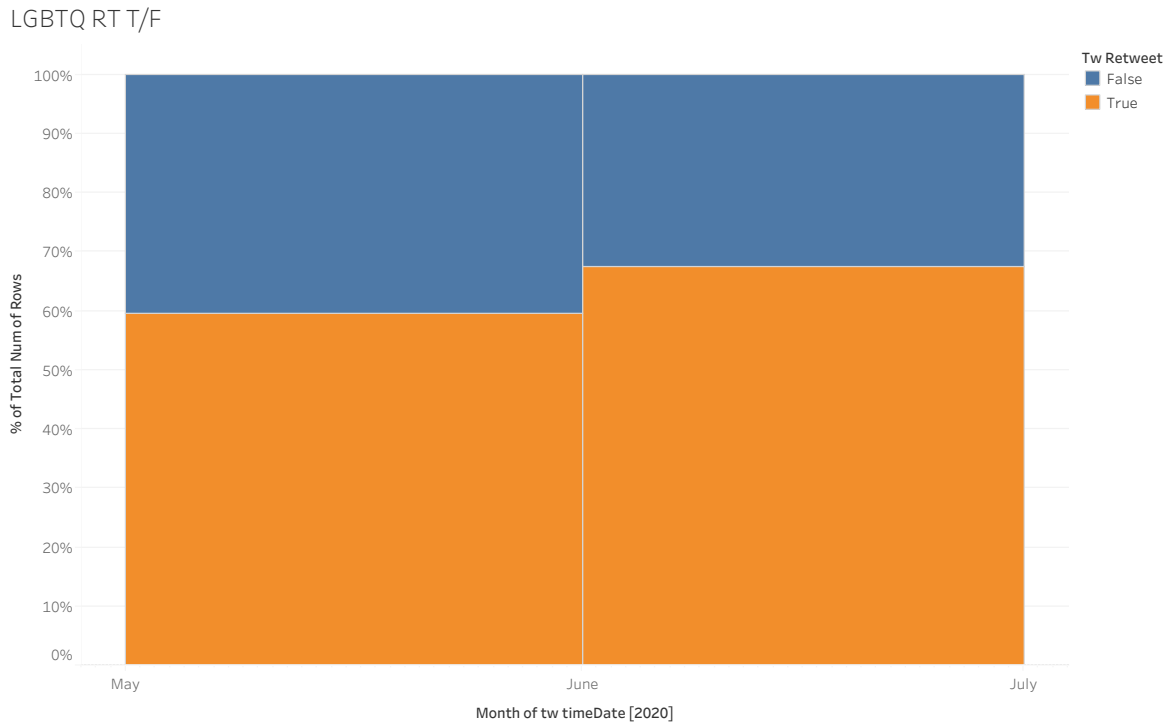
LGBTQ RT T/F

The plot of % of Total Num of Rows for tw timeDate Month. Color shows details about Tw Retweet. The data is filtered on tw timeDate Week, Search Term and tw timeDate Month. The tw timeDate Week filter has multiple members selected. The Search Term filter keeps 11 members. The tw timeDate Month filter keeps May and June. Percents are based on each column of each pane of the table.

Figure 6. Pride Month Percentage of Original Tweets

Figure 6 once again supports the Stieglitz & Dang Xuan (2013) conclusion that retweeting is a common response to Tweets that are emotionally charged. The Pride Month event includes many rallies, events and expressions of the community member's personal experiences. People both inside and outside of the community can retweet Tweets related to the aforementioned topics as a show of support.

In conclusion, Tweets relating to members of the LGBTQ+ community increased in number, average sentiment, and decreased in original tweet composition during the month of

June, commonly associated with "Pride Month". This indicated a positive effect on the sentiment relating to the LGBTQ+ community during this event.

Conclusion

Based on our results, we can conclude that U.S. national events do have impacts on people's sentiment of minority groups on Twitter. We saw correlations between each event and changes in the Tweet sentiment composition and the number of unique Tweets. Tweets and hashtags relevant to a certain people group are often used to organize masses of people and start a dialogue about an event (Theocharis, Lowe, van Deth, and García-Albacete, 2015). This indicates that once an event happens, a subgroup of the population (probably skewed to those who hold positions of power or influence) control the narrative, and therefore sentiment on that movement. People are more likely to Retweet their favorite politician or celebrity's opinion than draft their own on Twitter. Because we saw increased levels of sentiment during Pride Month but decreased levels in the Black Lives Matter Movement and Coronavirus Pandemic, we cannot state that there is a correlation between positive or negative sentiment and number of original Tweets. However, we can draw the conclusion that support is not necessary tied to a positive or negative average sentiment. Rather than positive sentiment solely indicating support, we find that a trend of sentiment leaning either way is an indicator of emotional valence people are having toward the greater conversation. For example, many people were very angry about the events that led up to and occurred during the Black Lives Matter Movement, yet still supported the movement. Tweets from people who felt this way would show a sharp negative sentiment, yet these people still supported the movement. The Black Lives Matter Movement and Coronavirus Pandemic both saw significantly decreased average sentiment composition. Both were controversial and politicized heavily during the

events. Because of historic levels of polarization in the United States currently, it can be presumed that when events are politicized, the average sentiment towards the people groups involved decrease. This correlation holds up when we look at the Pride Month event as well. As the LGBTQ+ movement has gained acceptance in American culture over the past 30 years, it has become less politicized and therefore even though there were fewer original tweets, average sentiment with regards to the community during the month went up.

REFERENCES

B. Suh, L. Hong, P. Pirolli and E. H. Chi, "Want to be Retweeted? Large Scale Analytics on Factors Impacting Retweet in Twitter Network," 2010 IEEE Second International Conference on Social Computing, Minneapolis, MN, USA, 2010, pp. 177-184, doi: 10.1109/SocialCom.2010.33.

C. Hutto and E. Gilbert, "VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text", ICWSM, vol. 8, no. 1, May 2014.

Dean, Brian, "How Many People Use Twitter in 2021". February 10, 2021. https://backlinko.com/twitter-users

Eunji Lee, Jung-Ah Lee, Jang Ho Moon, and Yongjun Sung.Cyberpsychology, Behavior, and Social Networking.Sep 2015.552-556.http://doi.org/10.1089/cyber.2015.0157

Stefan Stieglitz & Linh Dang-Xuan (2013) Emotions and Information Diffusion in Social Media—Sentiment of Microblogs and Sharing Behavior, Journal of Management Information Systems, 29:4, 217-248, DOI: 10.2753/MIS0742-1222290408

Sanjay Sharma "Black Twitter? Racial Hashtags, Networks and Contagion", New Formations, Number 78, Summer 2013, pp. 46-64(19), https://doi.org/10.3898/NewF.78.02.2013

Yannis Theocharis, Will Lowe, Jan W. van Deth & Gema García-Albacete (2015) Using Twitter to mobilize protest action: online mobilization patterns and action repertoires in the Occupy Wall Street, Indignados, and Aganaktismenoi movements, Information, Communication & Society, 18:2, 202-220, DOI: 10.1080/1369118X.2014.948035