An Association of Twitter Derived Measures of Discrimination and Race-Motivated Hate Crimes

Deja Dunlap[1] and Dr. Soo Park[1#]

[1]Advanced Technologies Academy
[#]Advisor

ABSTRACT

In 2020, there was a sharp rise in Anti-Asian American hate crimes. While the causes are still being explored, one of the main reasons this occurred was due to the rise in Anti-Asian American rhetoric connected to the COVID-19 pandemic. This exemplifies how prejudices often can turn into violence, more specifically hate crimes. And hateful speech, oftentimes spread through the internet, can help spread and deepen that prejudice to a wider audience. In previous research, prejuidice and discrimination has largely been studied through self-reported measures. While those studies are necessary, they are susceptible to many pitfalls and should be supplemented with statistics-based measures. This study employs one of these emerging measures. This study used the Twitter Streaming API to collect tweets containing a racial slur, an extension of hateful speech, geo-located from users in the United States over a period of two months. With these, this study compared statewide Twitter racial slur usage and statewide average sentiment polarity of the tweets with race-motivated hate crimes from the state. Though no relationship was found between statewide sentiment expressed and race-motivated hate crimes, this study did find a significant positive relationship between statewide Twitter slur usage and hate crimes. These findings support previous studies that have connected Internet-based measures of discrimination with aspects of general well-being. and even specifically hate crimes. This study shows the potential of using data derived from social media to create cost effective and expansive indicators of an area's social climate around a topic, even race.

Introduction

The Center for the Study of Hate and Extremism found that anti-Asian hate crimes went up by approximately 150% in 2020. While the reasons for the uptick are still being researched, Karthick Ramakrishnan credits a significant part of the blame to the upsurge of anti-Asian prejudice (Yam, 2021). Anti-Asian sentiments are not new (Usagi, 2021), but the 2020 wave of Anti-Asian rhetoric came from the weaponization of COVID-19's origins in China (Yam, 2021). The discriminatory language used against Asian Americans helped fostered an environment of hate that made actions likes hate crimes against Asian-Americans more acceptable, leaving the community more vulnerable to attacks and threats.

The increase in anti-Asian violence serves as an example of prejudice turning into violence. As Wiktor Soral ad his associates found, frequent exposure to hate speech, a form of discriminatory language, leads to desensitization to the speech and more prejudice towards the victims of the speech (Soral et al., 2018). This can have damaging effects on the communities being targeted by the hate speech, as observers will gradually be less able to identify it as hate speech and less willing to challenge it, thus allowing the discriminatory language to spread and intensify. This is dangerous because, as a study used by law enforcement to assess hate crimes found, nearly all motivations for hate crimes derive from some form of prejudice. By their very nature, hate crimes are crimes born of extreme prejudice (Levin & Devitt, 2008; American Psychological Association, 2017).

While hate crimes have gone down over the years, they remain a pervasive problem in the United States. In 2020, the Department of Justice (DOJ) found 8,052 cases of hate crimes involving 11,126 victims, with over 60% of those crimes being motivated by the victims' race, ethnicity, and/or ancestry (Department of Justice, 2020). These numbers don't include the vast numbers of crimes that go unreported for a variety of reasons (Usegi, 2021). These crimes can have damaging effects on the victims. Victims of these types of crimes are more likely to suffer from posttraumatic stress disorder, depression, anxiety, stress, and many other disorders when compared to victims of non-bias motivated crimes (American Psychological Association, 2017; Cramer et al., 2021). As Kate Usagi reported, it can even have effects on unaffected members of the community. Following the spike of violence and prejudice against Asian Americans in 2020, many of Usagi's own Asian American friends and family grew scared of leaving their home (Usagi, 2021). As the American Psychological Association (APA) explains, it's because members of the victim's group can feel a shared decreased sense of security and safety in their community. (American Psychological Association, 2021)

With this in consideration, researchers have found trying to measure prejudice challenging. Most research about prejudice is based on self-report measures, but these are susceptible to a variety of issues: participants lying, misleading questions, expectancy bias from researchers, along with a variety of other things that could affect results (Encyclopedia, n.d). While self-reported measures are important, they should be supplemented with more definitive statistics-based measures.

One of the novel forms of statistics-based measures is from the Internet, specifically from social media. Social media sites encourage the sharing of opinions and, because of the anonymity one can achieve on these apps, they can serve as outlets for prejudiced opinions. In fact, during 2020, social media was one of the mediums most often used to communicate discriminatory language about Asian Americans (Yam, 2021). This study uses data derived from Twitter to create an Internet-based measure of discrimination. Through this, this study may be able to create indicators of the social climate towards different racial groups in different regions of the United States.

Previous studies have used Internet-based measures to create indicators around a variety of topics. One study using data collected from Twitter looked at state-wide occurrences of talking about marijuana, finding a statistically significant positive relationship between the liberalness of laws surrounding the drug and the number of tweets referencing the drug (Daniulaityte et al., 2015). This study finding a significant relationship between references of a topic on social media and the environment that the users are generating them from establishes that Internet-based measures can provide an indicator for attitudes and the social climate about those topics.

While there is a limited amount of research utilizing social media to create indicators of discrimation specifically, there have been three studies finding associations between some effect on minorities and Internet-based discrimination. In one, David Chae and his associates examined the relationship between the proportion of Google searches containing the N-word and Black mortality, finding a significant positive relationship between them (Chae et al., 2018). Expanding on this, Thu Nguyen examined the relationship between sentiment expressed about minorities on Twitter and minorities women's birthing outcomes, finding that minority mothers in states where tweets contained less positive sentiments towards minorities were more likely to have higher rates of low birth weight and preterm birth (Nguyen et al, 2018). Nguyen later expands on this in a different study, where she specifically explores the relationship between sentiment expressed about minorities online and hate crimes and resident racial bias (Nguyen, et. al., 2021). Though she finds no significant relationship between the sentiment and hate crimes, she notes that this could be due to the fact that hate crimes often go unreported (Usegi, 2021). She was able to find a significant relationship between lower positive sentiment expressed about minorities on Twitter and higher rates of implicit and explicit racial bias, further showing the potential in this emerging methodology. While the previous studies could establish that social media can provide a useful resource when studying trends of a society, these studies establish social media can also be used to create measures of discrimination specifically.

This study focuses on racial slurs, or a derogatory term derived to a harm a particular group of people (Popa-Wyatt, 2018), because slurs are an extension of discriminatory language. As Mihaela Popa-Wyatt argues, slurs have the power to alter the dynamic of a conversation, creating an unjust power imbalance. She even argues that the use of these words helps spread and intensified oppression against the group the slur is directed towards (Popa-Wyatt, 2018). But the reclamation of many of these words by the groups they were originally directed towards cannot be ignored (Shenin & Thompson, 2014); Therefore, the context around the word being used will be accounted for through sentiment analysis.

The purpose of this study is use the Twitter Streaming Application Programming Interface (API) to examine the association of Twitter derived measures of discrimination and race-motivated hate crimes across the US. While previously mentioned studies have focused on the relationships of Internet-based discrimination and varying aspects of well-being, this study's primary focus is on a currently unexplored form of Internet-based discrimation - Twitter usage of slurs. This study uses Twitter data to create indicators of state racial climate by tracking the state-wide usage of racial slurs and the sentiment expressed around those slurs and compares those to statewide race-motivated hate crimes. I hypothesized that states with higher usage of racial slurs and lower positive sentiment towards racial and ethnic minorities will see higher rates of race-motivated hate crimes.

Methods

Selecting the Keywords

The keywords selected for the purposes of this study were derived from the OwlApps "List of Ethnic Slurs and Epithets by Ethnicity". The ones used in this study were those created to be directed to one of the racial/ethnic groups that the Federal Bureau of Investigation (FBI) tracks when collecting hate crimes reports - limiting the slurs to those directed towards people of White, African-American or Black, Arab, Asian, Hispanic or Latino, Native American or Native Alaskan, and Pacific Islander descendant (Federal Bureau of Investigation, n.d.). Slurs were excluded if they were deemed too general for it to be used in its derogatory meaning majority of the time. For example, the term "Jim Crow" can be used in a derogatory manner against an African American person, but is more regularly used to refer to segregation laws implemented in the US in the 19th and 20th century, so it was excluded from this list. When there were multiple variations of the same word, the most common iteration was selected. Slurs were also excluded if they didn't originate from or have regular usage in the US (OwlApps, n.d). With these qualifications, the list was narrowed down to 55 keywords (OwlApps, n.d.). Of that, 15 have historically been directed towards African American/Black people, 7 towards Native Americans, 6 towards Latino/Hispanic people, 17 towards Asian people, and 10 towards White people. In an effort to keep representation even, 5 keywords were then randomly selected from each racial group's list, leaving a total of 25 keywords that were used in this study. The full list of slurs tracked can be found in Appendix A.

Collecting the Tweets

This study focused on Twitter specifically because, with Twitter having over 300 million active users a month (Iqbal, 2022), and over 90% of users making their profiles public (Mislove, et. al., 2011), this study can reach a much wider portion of the nation than most self-reported studies. Furthermore, unlike social media sites like YouTube or TikTok, it is more centered on words than videos or photos, making it easier to track trends in national conversations. And, with its 280 character max per tweet, the information is much easier to conduct sentiment analysis on compared to other social media sites centered around longer posts like Reddit or Tumblr. Using the Twitter Streaming API, or an intermediary system that allows software applications to communicate with each (IBM, 2020), a StreamListener was set-up as a connection to Twitter where I tracked incoming tweets containing at least one of the keywords from when the program was first started to when it closed. Using the programming language Python, I developed an algorithm to filter the tweets based on certain parameters. All tweets that were not in English, didn't have a location attached to it, or originating from somewhere outside the US were excluded. Then, using the "place" attribute attached to the remaining tweets, the tweets were then geo-located to the state that they originated using the Nominatim Python package. While only about 1-2% of tweets are geo-tagged, where users attach their current locations when publishing a tweet, about 500 million tweets (Zote, 2021) are published everyday, leaving enough to create a representative sample. Once the tweet was deemed to have fit all of the criterias, the username of the tweeter, tweet content, state geo-located to the tweet, sentiment polarity score, and the keyword used in the tweet were then added to the database. The algorithm used to accomplish all of this can be found in Appendix B. Using the "Always-On" feature on the IDE Repl.it, the algorithm was left running continuously for a period of sixty days from January 1, 2022 to March 1, 2022. After the tweets were collected, they were filtered for duplicated tweets.

Sentiment Analysis

Due to the fact that many racial slurs have been reclaimed by certain members of the community that the slur was first directed toward, simply tracking the usage of the words wouldn't be enough to create indicators of discrimation for an area. That's why this study used sentiment analysis to track the polarity of a tweet. Sentiment analysis is a form of natural language processing that tries to extract the emotions tied to an expression (Es, 2021). Sentiment analysis was done for this project through the TextBlob Python package due to its simplicity and availability. As previously mentioned, before the tweet's were added to the database, they were given a sentiment polarity number, on a scale from 1.0 to -1.0, to track whether the tweet was (1.0) positive, (0) neutral or (-1.0) negative. For example, something like "I'm very happy" would receive a sentiment polarity score of 1.0, while "I'm very sad" would receive a score of -0.65. For clearer analysis, the tweets were fixed of any spelling errors before being passed through the sentiment analysis process.

Hate Crime Statistics

Every year, the FBI releases a report of the hate crimes reported during the previous year, breaking down the data by state and motivation. It must be noted that the crimes reported to the FBI are on a completely voluntary basis, meaning the police department chooses whether they want to share this information with the FBI (Federal Bureau of Investigation, 2021). Since the purpose of deriving tweets with racial slurs was to create indicators of racial discrimination, this study will be focusing on the race-motivated hate crimes only. This study used information derived from the latest report the FBI released in 2020.

Statistical Analysis

Because of the different natures of the measures of discrimination used in this study, two different analysis methods were used for the two independent variables - statewide racial slur usage and statewide average sentiment polarity. Nonetheless, all statistical analysis was conducted using the programming language, R.

When comparing the race-motivated hate crimes with Twitter racial slur usage by state, this study used negative binomial regression models. This was used in lieu of Poisson regression because data was over-dispersed. The model was then controlled for potential covariates, namely 2020 percentage of non-White population, 2020 population density (per square mile) (United States Census Bureau, 2021), and southern state indicator (yes/no) (Britannica, n.d.). The states were then divided into quantiles based on the state-wide Twitter slur usage, with the lowest quantile being used as the referent level. The coefficients from the model were then used to estimate incidence rate ratios (IRRs) for the relationship between the two variables of interest.

When comparing the statewide average sentiment polarity score and race-motivated hate crimes, negative binomial regression modeling could not be employed. While the data was still over-dispersed, there were negative values included in the dataset, making it ineligible for that form of regression modeling. Instead, linear modeling was used, controlling for the same potential covariates. The states were then divided into quantiles based on the state-wide average sentiment polarity scores, with the lowest quantile being used as the referent level. The coefficients from the linear modeling were then used to estimate IRRs.

Ethics Statement

This study involved secondary analysis of existing data and therefore is exempt from human subjects review. Hate crimes statistics were collected from the Federal Bureau of Intelligence public records, which are anonymized and publicly available. Identifying information about twitter users, specifically their usernames, were collected only to identify the number of unique users in the dataset and were deleted from the dataset after it was determined.

Results

Descriptive Characteristics of the Tweets

Table 1. Descriptive Characteristics of Twitter Sentiment Polarity and Racial Slur Usage

Descriptive Characteristics of the Independent Variables		
	Sentiment Polarity	Racial Slur Usage
Average	0.014	849.961
Median	0.006	302
Variance	0.038	2034096.918

After a period of sixty day, 43338 geotagged tweets containing one or more of the keywords were collected from 12762 unique users. Of those, 42539 contained a slur predominantly directed towards African Americans, 334 contained a slur directed towards Hispanic people, 314 directed towards White people, 148 directed towards Native Americans or Pacific Islanders, and 222 directed towards Asian Americans. The very clear unequal skew towards slurs directed towards African Americans is most likely because of the unparalleled popularity of the “N-word”. Single handedly, the “N-word” accounted for over 97% of all tweets collected, as shown in the table in Appendix C. Table 1 highlights the descriptive data of the state-level Twitter-usages of racial slurs and the Twitter-derived indicators of sentiment towards racial/ethinc groups. As previously mentioned, the mean and variance for both measures were drastically different. And while a few drastic outliers upshot the mean count for tweets containing one of the slurs, most states saw counts much closer to the median. For the sentiment polarity, most states hovered around neutral territory when tweeting with a racial slur.

Geographical Distribution of Tweets

Figure 1 illustrates the geographic distribution of racial slur usage in tweets. The states with the highest usage were Texas (7039), Georgia (5228), California (4718), New York (3235), and Florida (2693). The states with the lower slur usage were South Dakota (6), North Dakota (8), Vermont (13), Montana (13), and Wyoming (15). The states with the least positive sentiments when using racial slur were North Dakota (-0.079), Wyoming (-0.053), Connecticut (-0.042), Oregon (-0.028), and Arkansas (-0.013). The states with the most positive sentiment when using racial slur were West Virginia (0.128), Montana (0.113), Idaho (0.093), Vermont (0.080), and Maine (0.074). The complete order list of racial slur usage and average sentiment polarity can be found in Appendix D.

Table 2. Negative Binomial Modeling Estimating Incident Rate Ratios for Race-Motivated Hate Crimes

	Incident Rate Ratios (95% CI)	P-value
Quantile 2	4.447 (1.990, 10.296)	=.001***
Quantile 3	6.617 (2.907, 15.253)	<.001***
Quantile 4 (highest)	11.161 (4.255, 30.389)	<.001***
Non-White Population	0.998 (0.97, 1.02)	0.86
Population Density	0.910 (0.91, 1.00)	0.79
Southern State Indicator (No)	2.10 (0.20, 9.97)	0.42
Southern State Indicator (Yes)	1.17 (0.11, 6.176)	0.87
Akaike Information Criteria (AIC)	560.34	

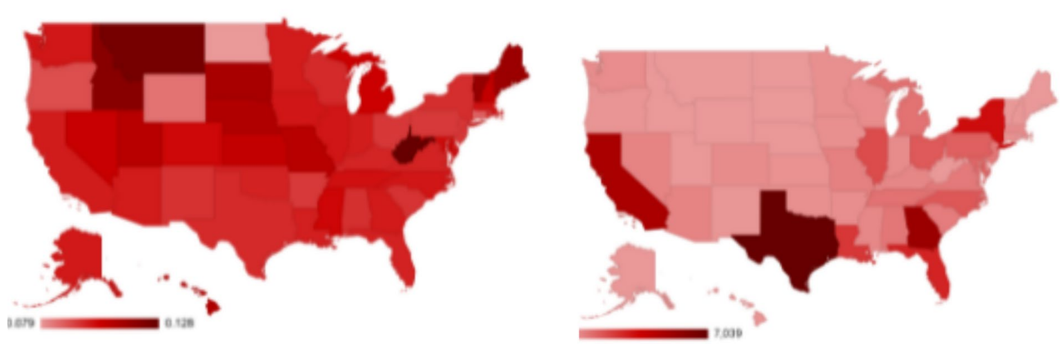

Figure 1. Graph on the right illustrates the geographical distribution of sentiment polarity of tweets by state, graph on the left illustrates the geographic distribution of slur usage in tweets by state

Associations between Racial Slur Usage and Hate Crimes

Results from negative binomial regression associating state-wide slur usage with hate crimes, controlling for potential covariates are shown in Table 2. I found a significant positive relationship between the two variables. When compared to states in the first quartile, states in the second quartile saw an IRR of 4.48 (95% CI: 1.98,10.30, p=0.001), states in the third quartile saw an IRR of 6.17 (95% CI: 2.91, 15.25, p<0.001), and states in the fourth quartile saw an IRR of 11.16 (95% CI: 4.26, 30.39, p<0.001). The wide range in the confidence intervals was likely due to the relative rarity of hate crimes. Interestingly though, the potential covariates included in the model - state non-White population, population density, and southern state indicator - had seemingly no effect on the model at all. Removing them left the result identical to the model with them.

Associations between Sentiment and Hate Crimes

Results from linear modeling associate average state sentiment polarity scores with hate crimes, controlling for potential covariates, are shown in Table 3. I found no significant relationship between average sentiment polarity and state hate crimes, with the IRRs for each of the groups being too low and the confidence interval too wide to have any meaning. Likewise, the potential covariates continued to have little effect on the model.

Table 3. Simple Linear Modeling Estimating Incident Rate Ratios for Race-Motivated Hate Crimes

	Incident Rate Ratio (95% CI)	P
Quantile 2	1.83e^29 (9.28e^-20, 3.64e^87)	0.202
Quantile 3	6.05e^-10 (3.6e^-54, 1.01e^35)	0.676
Quantile 4 (highest)	2.05e^-13 (8.739e^-144, 4.84e^177)	0.845
Non White Population	2.94e^+1 (1.21,7.10e^02)	0.0379*
Population Density	9.87e^-1 (9.61e^-1,1.01)	0.332
Southern State Indicator (No)	6.58e^28 (7.42e^101, 5.84e^157)	0.654
Southern State Indicator (Yes)	3.05e^-10 (3.00e^144, 3.10e^124)	0.887

Discussion

The descriptive characteristics of the tweets are interesting in how they fit in with the current literature. In regards to the keyword composition of the tweets, it may initially seem surprising, but, when looking at other studies that have employed a similar method, it’s clear that it agrees with at least some of the previous literature. For example, in Nguyen's study focusing on minority women’s birthing outcomes, she tracked tweets containing one or more of nearly 400 race-related keywords (over 15 times the amount this study tracked), and the “N-word” still single-handedly showed up in nearly half of all tweets collected (Nguyen, et. al, 2018). In accordance with this, it seems that future researchers hoping to employ this method will have to take into account the unparalleled popularity of the "N-word" and decide whether the word should be included in the keyword database based on the purpose of their study.

This study's results would suggest a strong positive relationship between statewide Twitter racial slur usage and race-motivated hate crimes. This could be because social media serves as an outlet to voice opinions, no matter how prejudiced they may be, as seen by the previous studies that have used Internet derived measures to create social climates around different topics (Chae, et. al, 2018; Nguyen, et.al, 2021; Nguyen, et. al, 2018; Daniulaityte et al., 2015; Nguyen, et. al, 2017). This would agree with previous studies assertions that the discriminatory language and attitudes we see on Twitter would then translate into discriminatory behaviours in the real world. And, as previously stated, hate crimes are crimes born of discrimination (Levin & Devitt, 2008; American Psychological Association, 2017).

This study also saw no significant relationship between average sentiment polarity and hate crimes. This could be due to the reclamation of many racial slurs previously mentioned. As seen in **Table 1**, most states' sentiment polarity scores were around zero, largely neutral in nature. Because of the changing connotations of these words, these words are increasingly being used in a less negative light. While the initial connotations of the words can never truly be separated from the words themselves, they can grow more distant from each other (Shenin & Thompson, 2014). This study does differ from previous studies in this fact. While Nguyen's study focused on comparing Twitter-derived sentiments about minority-to-minority women's birthing outcomes, she did find a significant relationship between them. In turn, she found a significant relationship between sentiment expressed about minorities and some effect of that discrimination. But that doesn't necessarily mean that our studies contradict. The sentiment analysis tool that Nguyen employed in their study was much more advanced than the one used here. And, while the sentiment analysis tool I used utilized a scale that differentiated between negative (-1.0), neutral (0), and positive (1.0) sentiment, Nugyen's scale focused on positive (1.0) and not positive (0) scoring. But, in general, sentiment analysis itself is a relatively new tool, so future research could be done on how to best adapt sentiment analysis tools for analyzing online conversations.

Conclusion

It was initially hypothesized that states with higher Twitter racial slur usage and lower positive sentiment towards racial and ethnic minorities will see higher rates of race-motivated hate crimes. And this did partially come true. States in the highest, second highest, and third highest quantile for racial slur usage were 4.48, 6.17, and 11.16 times more likely to see hate crimes compared to states in the lowest quantile. Though no significant relationship was found between statewide average sentiment and race-motivated hate crimes, that could be because of the changing connotations of many slurs.

This study did face some limitations. Firstly, this study faced a much smaller sample size and time frame than other studies that have employed this research method. Normally, the research team will collect tweets for a period of several months to even years, collecting hundreds of thousands of qualifying tweets (Chae, 2015; Danuilayte, 2015; Nguyen, et. al. 2017; Nguyen, 2018). Additionally, the sentiment analysis tool that I relied on for this study is relatively simple. Even the best of modern sentiment analysis tools aren't able to properly analyze things like humor or sarcasm, affecting this study's indicators of sentiment towards racial groups in an area. The results could have also been affected by comparing outcomes from 2020 to tweets that originated in 2022. As far as the geo-location, that stands the risk of users lying about their current location. It's also unable to differentiate users that are visiting a certain location rather than living there. Additionally, users of social media tend to be younger than the general population; in 2016, 36% of individuals aged 18–29 years old used Twitter compared to 21% of individuals 50–64 years, and 10% among those 65 + years (Greenwood, Perrin, & Duggan, 2016). But usage of social media has been steadily increasing.

Lastly, the research method employed in this study, just like other research methods, relies on people's willingness to share information. While the internet may act as a way to vent personal opinions about certain topics, the fact that that information will forever be linked to them may keep some users from fully expressing their opinions in such a public manner.

Despite these limitations, this study may have important implications for methods to measure and monitor racism.This study is one of a few that uses the relatively new technology of social media to create indicators of social climate around an issue. It's one of even fewer to create indicators of discrimination from them. This could have huge implications for the future of discrimination research. Where as, most previous discrimination research focuses on self-report measures, this could open the door to further statistic-based research that could supplement the previous established body of research. This method is also much more cost-effective and widespanding than most self-reported methods, allowing for better generalizations to the population. This study results suggest that discriminatory attitudes expressed on social media can translate into discriminatory behaviour in the real world. This could be useful to law enforcement agencies looking to predict and curtail race-motivated hate crimes in areas disproportionately affected by it.

Acknowledgements

Special thanks to Dr. Soo Park, Kirsten Taylor, and Dr. Anthony Marentic for their help with this research project.

References

Britannica. (n.d.). *the South | Definition, States, & History | Britannica*. Encyclopedia Britannica. Retrieved March 28, 2022, from https://www.britannica.com/place/the-South-region

Es, S. (2021, November 12). *Sentiment Analysis in Python: TextBlob vs Vader Sentiment vs Flair vs Building It From Scratch - neptune.ai*. Neptune.ai. Retrieved November 30, 2021, from https://neptune.ai/blog/sentiment-analysis-python-textblob-vs-vader-vs-flair

FBI. (n.d.). *Hate Crime Statistics*. FBI. Retrieved November 30, 2021, from https://www.fbi.gov/services/cjis/ucr/hate-crime

OwlApps. (n.d.). *List of ethnic slurs by ethnicity*. owlapps. Retrieved November 30, 2021, from http://www.owlapps.net/owlapps_apps/articles?id=8308199&lang=en

Zote, J. (2021, March 11). *12 Essential Twitter Stats to Guide Your Strategy*. Sprout Social. Retrieved November 30, 2021, from https://sproutsocial.com/insights/twitter-statistics/

Chae, D. (2015, April 24). *Association between an Internet-Based Measure of Area Racism and Black Mortality*. PLOS. Retrieved January 25, 2022, from https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0122963

Cramer, R., Fording, R., & Gerstenfeld, P. (n.d.). *Hate-Motivated Behavior: Impacts, Risk Factors, And Interventions*. HealthAffairs. https://www.healthaffairs.org/do/10.1377/hpb20200929.601434/full/

Daniulaityte, R., Nahhas, R., & Wijeratne, S. (2015, October 1). *"Time for dabs": Analyzing Twitter data on marijuana concentrates across the US*. PubMed. Retrieved January 25, 2022, from https://pubmed.ncbi.nlm.nih.gov/26338481/

Department of Justice. (2020). *Hate Crime Statistics | HATE CRIMES*. Department of Justice. Retrieved January 25, 2022, from https://www.justice.gov/hatecrimes/hate-crime-statistics

Encyclopedia. (n.d.). *Self-Report Method*. Encyclopedia.com. Retrieved January 25, 2022, from https://www.encyclopedia.com/social-sciences/applied-and-social-sciences-magazines/self-report-method

FBI. (n.d.). *Hate Crime Statistics*. FBI. Retrieved November 30, 2021, from https://www.fbi.gov/services/cjis/ucr/hate-crime

Greenwood, S. Perrin, A., Duggan, M. (2016) *Social media update 2016.* Pew Research Center, Washington, DC (2016) http://downtowndubuque.org/wp-content/uploads/2017/01/Social-Media-Update-2016.pdf

IBM. (2020, August 19). *What is an Application Programming Interface (API)*? IBM. Retrieved January 25, 2022, from https://www.ibm.com/cloud/learn/api

Iqbal, M. (2022). *Twitter Revenue and Usage Statistics (2022)*. Business of Apps. Retrieved January 25, 2022, from https://www.businessofapps.com/data/twitter-statistics/

Krieger, N, Waterman, PD, Kosheleva, A, Chen, JT, Carney, DR, Smith, KW, et al. (2008) *Exposing racial discrimination: implicit & explicit measures - the My Body, My Story study of 1005 US born black & white community health center members*. PLoS One

Levin, J., & Devitt, J. (2008). *Hate Crimes*. Jack Levin. Retrieved January 25, 2022, from https://jacklevinonviolence.com/articles/HateCrimesencyc92206FINAL.pdf

Nguyen, Q.C., Meng, H., & Li, D. (2017). Social media indicators of the food environment and state health outcomes. *Public Health*, *148*, 120-128. doi.org/10.1016/j.puhe.2017.03.013

Nguyen, T. (2018). Associations with Birth Rate and Preterm Birth in the United States. *Computers in Human Behavior*, *89*. EBSCOhost. doi: 10.1016/j.chb.2018.08.010

Nguyen, T., Huang, D., Michaels, E., Glymour, G., Allen, A., Nguyen, Q., (2021). Evaluating Associations Between Area-Level Twitter Expressed Negative Racial Sentiment, Hate Crimes, and Residents' Racial Prejudice in the United States. *SSM - Population Health*, *14*. doi: 10.1016/j.ssmph.2021.100750

Popa-Wyatt, M. (2018). Slurs, Roles, and Power. *Philosophical Studies*, *175*(11). EBSCOhost. doi: 10.1007/s11098-017-0986-2

Puente, A. E. (2017). *The Psychology of Hate Crimes*. American Psychological Association. Retrieved January 25, 2022, from https://www.apa.org/advocacy/interpersonal-violence/hate-crimes

Shenin, D., & Thompson, K. (2014). *Redefining the Word: Examining a Racial Slur Entrenched in American Vernacular That is More Prevalent than Ever*. The Washington Post. https://www.washingtonpost.com/sf/national/2014/11/09/the-n-word-an-entrenched-racial-slur-now-more-prevalent-than-ever/

Soral, W., Bilewicz, M., & Winiewski, M. (2018). Exposure to hate speech increases prejudice through desensitization. *Aggressive Behavior*, *44*(2), 136-146. EBSCOhost. 10.1002/ab.21737

United States Census Bureau. (2021, August 12). *Race and Ethnicity in the United States: 2010 Census and 2020 Census*. Census Bureau. Retrieved March 28, 2022, from https://www.census.gov/library/visualizations/interactive/race-and-ethnicity-in-the-united-state-2010-and-2020-census.html

Usagi, K. (2015). Confronting the Surge in Anti-Asian American Violence. *Humanist*, 155. doi.org/10.1016/j.drugalcdep.2015.07

Yam, K. (2021, March 9). *Anti-Asian hate crimes increased by nearly 150% in 2020, mostly in NY and LA, new report says*. NBC News. Retrieved November 26, 2021, from https://www.nbcnews.com/news/asian-america/anti-asian-hate-crimes-increased-nearly-150-2020-mostly-n-n1260264