The Resurgence of Cyber-Racism COVID-19 Pandemic and its After Effects

Akash Dutt Dubey

Submitted to: JMIR Public Health and Surveillance
on: May 03, 2020

Table of Contents

The Resurgence of Cyber-Racism COVID-19 Pandemic and its After Effects

Akash Dutt Dubey[1] PhD

[1]Jaipuria Institute of Management, Jaipur Jaipur IN

Corresponding Author:
Akash Dutt Dubey PhD
Jaipuria Institute of Management, Jaipur
Pratap Nagar, Sanganer
Jaipur
IN

Abstract

Background: With the increasing number of patients of COVID-19 all over the world, a particular section of the world population has been blaming China and World Health Organization for the spread of this disease. Due to this, a number of cases related to racism and hatred have been encountered all over. With Donald Trump using the term Chinese Virus, this cause has gained momentum and Ethnic Asians are now being targeted. The online situation also looks the same where certain segment of people are posting hateful comments and posts.

Objective: The research study has been done to analyze the increasing racism online during COVID-19. This research aims to analyze the emotions and sentiments that have been associated with the terms like CHinese Virus, Wuhan Virus and Chinese Corona Virus.

Methods: 16,000 tweets ranging from 11th April 2020 to 16th April 2020 were analyzed to find out the sentiments and emotions associated with the posts. This research has been done on R. Twitter API and SentimentR package were used for the collection of the tweets and then evaluating the sentiments respectively.

Results: This results suggested that majority of the tweets were of negative sentiments and carried emotions of fear, sadness, anger and disgust. There were a high usage of slurs and profane words. Also, terms like China LieD People Died, Wuhan Health Organization, Kung Flu, China Must Pay and CCP is Terrorist were used a lot in these tweets.

Conclusions: This study gives an insight of the rising hatred and cyber racism on Twitter. This study has analyzed the twitter posts and concludes that there are a majority of people who are tweeting with mostly negative sentiments for ethnic Asian, China and WHO.

(JMIR Preprints 03/05/2020:19833)
DOI: https://doi.org/10.2196/preprints.19833

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Introduction

Ever since its evolution, the social networks have always given a platform to the people worldwide to express their views and opinions. While in 1993, New Yorker [1] had explicitly stated in a cartoon that no one is a dog on internet (thereby signifying that the caste, race, ethnicity, religion and appearance does not matter when you are on internet), Nakamura [2] denied any existence of this utopian model and suggested that the Internet is “an outstanding example of racist medium”. Brown [3] in his research has concluded that the internet has often been a place where racism is disseminated in the various ways, including certain websites. These websites used offensive stereotypes to set the white supremacy over ethnic Africans. In 2011, Clark et al. [4] analyzed the weblogs using modified consensual qualitative research to study the different types of racial micro aggression which targeted American Indians. There have been sufficient studies that verify that there has always been racial aggression and hatred for different races, ethnicities and religions on internet.

In the current scenario, the whole world is facing the brunt of COVID-19 or nCoV19. The starting of COVID19 is said to be in the starting of December 2019, when several patients from Wuhan, Hubei Province reported severe respiratory infections. These patients had a background of working in the wholesale fish and seafood market, also known as wet markets [5]. In January 2020, the markets were completely closed down and disinfectants were used to sanitize them. On 7th January 2020, the researchers isolated a novel coronavirus which was referred as SARS-CoV-2 or 2019-nCov. Initially the World Health Organization denied the possibilities of human-to-human transmission 2019-nCoV on 11th January 2020. However, the confirmed cases continued to soar and on 30th January 2020, World Health Organization declared this COVID19 a Public Health Emergency of International Concern (PHEIC) and an epidemic. Finally, on 11th March 2020, WHO declared COVID-19 as a pandemic, and since there has not been any specific treatment to this disease, WHO recommended isolation and lockdown to reduce the spread of COVID-19.

On 17th March 2020, 45th President of United States of America, Donald Trump posted the following tweet: “*The United States will be powerfully supporting those industries, like Airlines and others, that are particularly affected by the Chinese Virus. We will be stronger than ever before!*” [6]. Instead of using the word Corona Virus, the word Chinese Virus sparked a series of controversies and people who were in support of Donald Trump started trending #ChineseVirus and #WuhanVirus [7, 8, 9] on various online platforms for social networking, Twitter being the most prominent of them. Racial slurs and profane words for Asian communities have been visible on Twitter ever since [10][11]. In Italy, there have been several incidents reported which dealt with anti-Chinese racism and discrimination. It is also believed that the increasing rate of xenophobia in Italy was the result of the

circulation of the information related to racism. [12]. In a previous work done by Budhwani and Sun[13], they stated that there has been 10x average increase in the usage of the words China Virus or Chinese Virus.

This research work has been conducted keeping in mind that there has been increase in the cyber racism and online hatred in the COVID-19 phase. The main aim of this research is to analyze the sentiments and emotions associated with the tweets which mention Chinese Virus or Wuhan Virus. This research also analyzed the most frequently used words while posting these tweets.

METHODOLOGY

Perceived as one of the most popular microblogging service provider, Twitter was launched in 2006. Initally, Twitter used to provide a limit of 140 characters which was later upgraded to 280 characters. Estimated number of Twitter users is 330 million worldwide. Twitter has been often used as a platform where people disseminate information, as well as share their opinion and emotions. This rapid sharing of the opinions of the user on Twitter has encouraged the researchers to determine the sentiments on almost everything, including sentiments towards products, movies, politics, digital technology and natural calamities [14, 15, 16, 17, 18].

Sentiment analysis of the twitter has also been used to derive the sentiments of the population regarding different diseases. Previously, sentiment analysis of twitter posts has been done to study the topic coverage and sentiments regarding Ebola virus [19]. This study analyzed two media sources separately i.e. Twitter and News Sources. Similarly, a study was conducted to study the key topics which influenced the negative sentiments on Twitter regarding Zika virus [20]. Sentiment analysis of the tweets was also done to analyze the sentiments of the patients who were affected by Crohn's disease. This study was done to gain the patients' perspective on a specific medical therapy [21].

While there is no single accepted psychological theory of basic human emotions, most of the studies accept the theory that a simple positive-negative dichotomy cannot be used to decide the human emotions as a whole. On the same lines, it is believed that the automatic sentiment analysis must also implement finely tuned algorithms to detail the human emotions. SentimentR is one such package which tries to evaluate the sentiments and emotions associated with texts [22]. Sentimentr package has been successfully used in analyzing the sentiments of the tweets which dealt with migraine activity [23]. It has been also used to analyze the tweets of Donald Trump and analyzed the relation between the sentiments of the tweets and the number of retweets [24]. In a review of four different sentiment computation packages, Maurizio Naldi concluded that the critical issue of negators is accurately dealt in the SentimentR package [25]. In other words, SentimentR was accurate is

calculating the difference between words like “useful”, “not useful”(negator), “really useful” (amplifier) and “hardly useful”(deamplifier). The potential of this package to calculate the sentiments based on the role of negators, amplifiers and deamplifiers was the reason this package was used to analyze the sentiments of the tweets in this study.

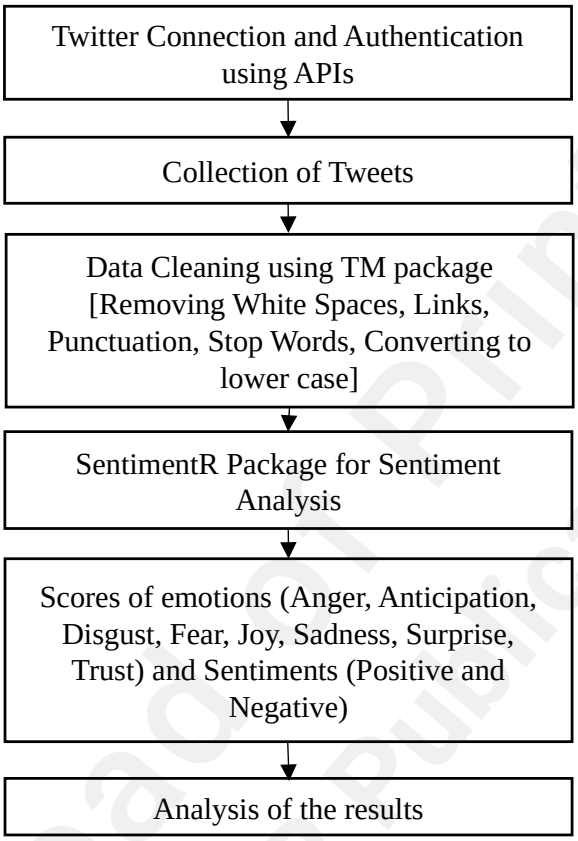


Figure 1 Flowchart for Sentiment Analysis of the tweets

Figure 1 illustrates the flowchart for the sentiment and emotion analysis for the tweets. The tweets were collected by using RTweet package in R.

For the collection of the tweets, search_tweets function of Rtweet package was used. In the collection process, the keywords to fetch the tweets were #ChineseVirus, #ChineseVirusCorona and #WuhanVirus. The date range of the searching was fixed from 11th April 2020 to 16th Aprilm 2020. The search process did not collect the retweets and replies, so that the duplication of the data can be avoided.

After the tweets were collected, the data cleaning process was performed using Text Mining (TM) package in R. This package was used to remove the white spaces, punctuation, stop words were removed and the tweets were converted to lower case. Post data cleaning, the SentimentR package was applied to analyze the tweets. Once the scoring of the tweets was done on the basis of sentiments

and emotions, the terms related to positive and negative sentiments, profanity and emotions were also calculated for further analysis.

Results

Using the tweet collection process, total 16,000 tweets were collected for the analysis. The collected tweets were analyzed using sentimentr package and the scoring was done on the basis of positive and negative sentiments. The SentimentR package in R scores the sentiments on a scale where 0 is considered as neutral, negative numbers indicate presence of negative sentiments and positive numbers indicate presence of positive sentiments. The sentiment score of each tweet was calculated individually and then the complete report of the sentiment across all the tweets was generated. The distribution summary of the sentiment analysis of the tweets has been shown in the table 1. The minimum value obtained in the analysis is -1.93076, which is the score of the tweet with most negative sentiment. The maximum score obtained during the analysis is 5.37143 is the score of the most positive tweet. The mean and median values of all the tweets were also calculated in the analysis. In table 1, the median and mean of the sentiments are -0.01581 and -0.06288 respectively. This shows that the sentiments observed in the tweets are skewed on the negative side, which shows that that the number of tweets with negative sentiments were more than the number of positive sentiments.

Table 1 Distribution Summry of Sentiment Analysis of the Tweets

Min	Median	Mean	Max
-1.93076	-0.01581	-0.06288	5.37143

Figure 2 shows the emotion analysis of the collected tweets. While the sentiment analysis of the tweets provide an overview of how the people were tweeting, the emotion analysis gives an insight of why this was happening. The results of the emotion analysis of the tweets has been illustrated in figure 3. It can be seen in the figure that the tweets with fear quotient is almost equal to the tweets with the trust quotient. When the four negative emotions (fear, sadness, anger and disgust) were analyzed collectively, the total percentage of these four emotions turned out to be 52.18%. While this result reconfirms the presence of majorly negative sentiments in the tweets, it also answers the constituents of the negative sentiments in the tweets. Three sample tweets of each emotion has been given in the table 2.

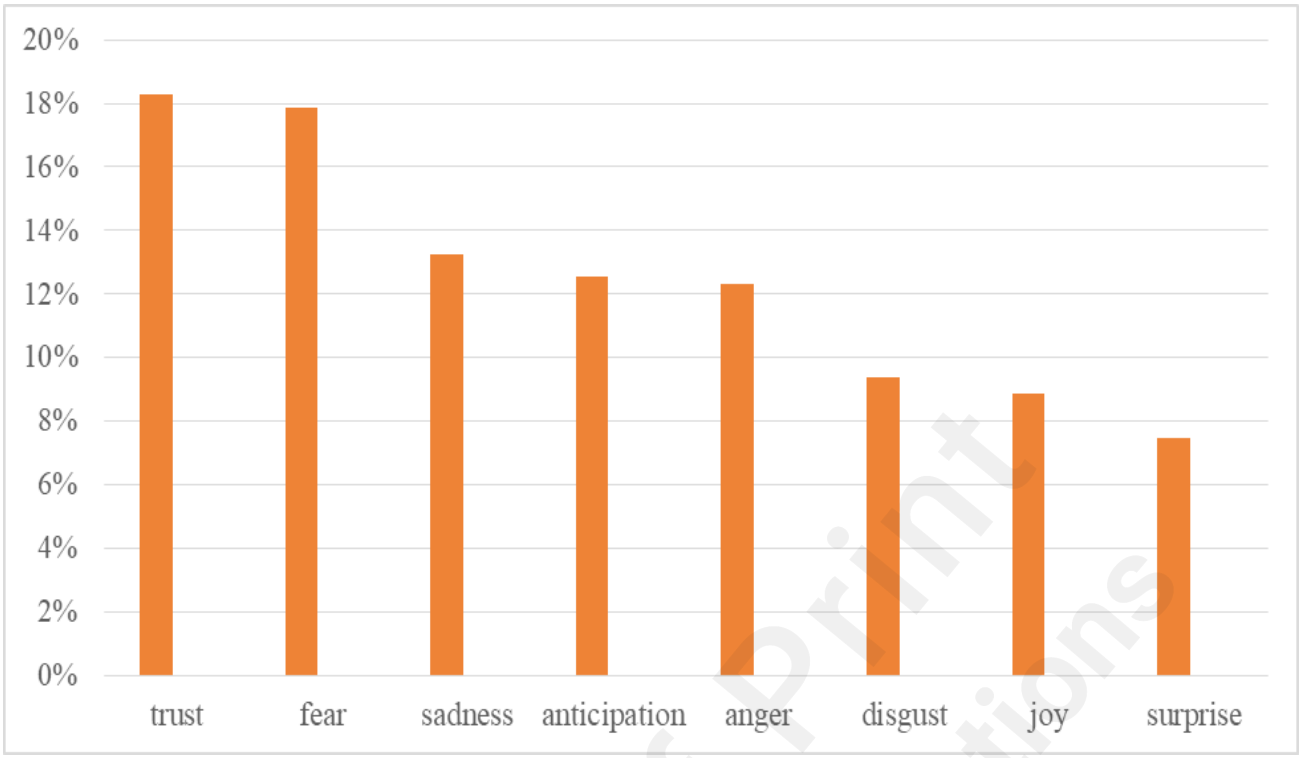


Figure 2 Emotion Analysis of Tweets

Table 2 Sample Tweets for Different Emotions

Trust	*"CMgovt has done very good job to handle WuhanVirus problem in state"* *"Breaking News US Government gave Wuhan Virology lab aMillion grant for virus research* *CoronaVirusUpdate Coronavirus COVID Covid WuhanVirus BreakingNews"* *"Great now Biden needs to explain that him calling the Chinese travel ban xenophobic would NEVER happen again and he made a HUGE mistake cause this attitude would kill us all if he was president ChineseVirusCorona"*
Anger	*"Good for Trump Calling the virus what it is the Wuhan virus Dont know why it triggers you so badly but you always have to find something daily to bang Trump around It gets old and you look childish WuhanVirus WuhanFlu"* *"Chinas carelessness and deceit has crashed global economies and costed countries trillions of dollars Every cent of debt that they hold from other countries should be forgiven Coronavirus WuhanVirus ChinaLiedPeopleDied"* *"I guess Twitter should ban all the Chinese because its ban in China ChinaLiedPeopleDied ChineseVirus"*
Sadness	*"Sir we had great hopes with you but it all shattered into pieces You r siding with the evil of CCP in this difficult timeshas become a mouthpiece for communist chinafunding it would be a crime against humanity ChineseVirus"* *"Just look at the economic destruction the WuhanVirus has inflicted on the American ppl"*

	"ChineseVirus COVID We lost th member of familyfriend in the CoronaWarfound ve hospitalised Yesterday world heard news of a mths old corona infected baby I plead to my friendsfollowers Corona is a seriousreal KeepSocialDistancing StayHomeSaveLives"
Anticipation	*"Considering most of the vaccines across globe come from India India might play a vital role in discovery of vaccine on ChineseVirus VaccinesSaveLives"* *"Our life is going to be changed drasticallyBe prepared for itChineseVirus"* *"I KNEW IT I think all countries should declare war on China WHO ChinaLiedPeopleDie Wuhan ChineseVirus ChinaMustPay ChinaVsTheWorld"*
Fear	*"We are heading for a long haul I am afraid in this fight against ChineseVirus with this pandering to a community which was not supposed to be in India since"* *"We are afraid of ChineseVirus so we are retreating now"* *"Sorry i afraid of ChineseVirusCorona ChineseVirus"*
Disgust	*"Will you be so shameless to buy Chinese mobile phones after ChinaVirus ChineseVirusCorona coronavirusindia coronavirus COVID covidindia Hit China hard where it will pain them most PS A branded phone fetches Chinese company more profit than selling components"* *"Shame on you ChineseVirusCorona"* *"Shameful the very same people who caused the WuhanVirus pandemic is now discriminating against innocent Africans"*
Joy	*"Happy Thai New Year Buddy Lets fight together MilkTeaAlliance FightForFreedom StandwithHK hkisnotchina TaiwanIsaCountry nnevvyy ChineseVirus ChinaMustPay ChinaLiedAndPeopleDied"* *"Not only recovered but got raised big fundLovely D coronavirus ChineseVirus WuhanVirus"* *"I get enough to live comfortably This ChineseVirus just depleted my savings Im happy with my investment"*
Surprise	*"Am surprised we still trust China havent we learnt our lesson ChineseVirus"* *"Shocking Did you know your taxwere being spent on this So is NIH partially responsible for WuhanVirus"* *"The world is still in utter shock . Right from the start experts advised the president to refrain from labelling COVIDa ChineseVirus to no avail It appears whatever DonaldJTrump was harboring against China has finally started manifesting in life threatening developments"*

While analyzing the tweets, 15 most frequent words which emoted different emotions were also analyzed. In figure 2 the results of the analysis has been illustrated. Words like death, good, money, pay, pandemic, trump, organization were most frequently used by the people while mentioning terms like ChineseVirus, WuhanVirus and ChineseVirusCorona. The presence of words like Death, Pay, Pandemic, Evil, Disease were repeatedly used in the tweets associated with the negative sentiments and emotions. These results, combined with the figure 2 and table 1, reflect the negative sentiments and emotions that have been communicated online.

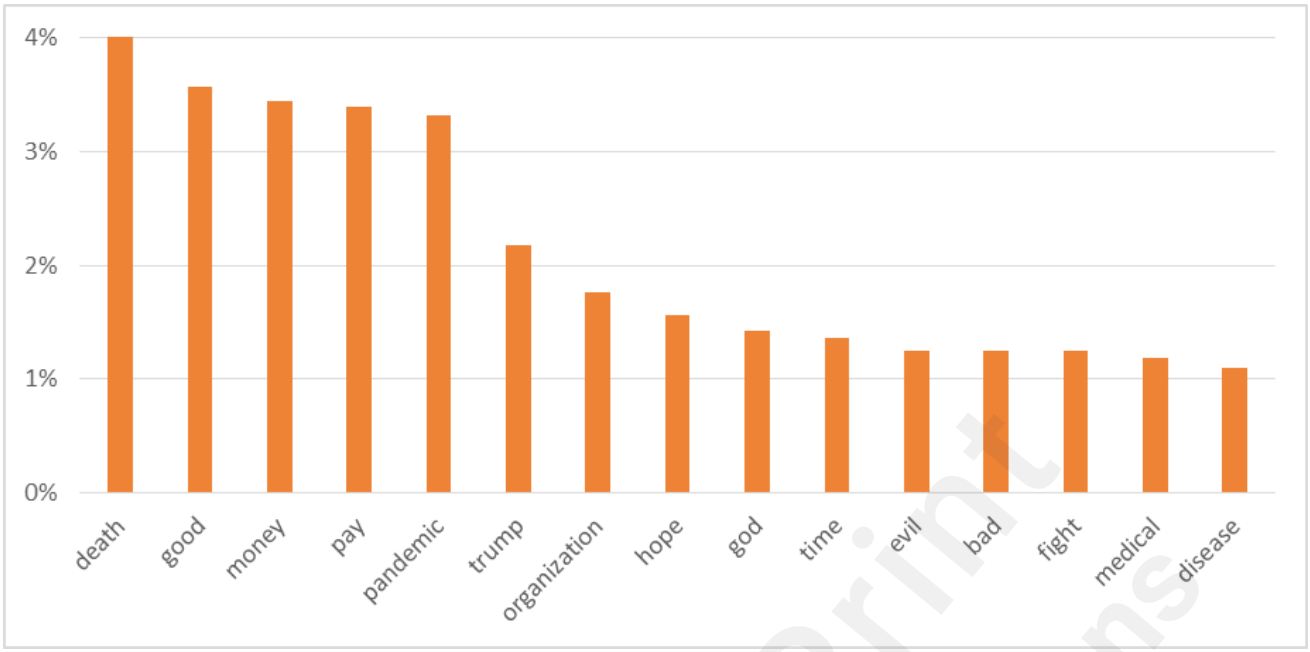


Figure 3 Frequency of most used words in the analyzed tweets

Discussions

From the results obtained in the analysis, the negative sentiments and emotions associated with the collected tweets were evident. A good number of tweets done with the term Chinese Virus had signs of hatred, disgust, fear and anger. Apart from the words which were associated with different emotions, there were some slangs created by the users which were not detected by the sentimentr package. Prominent among those were ccpisterrorist, ccpliedpeopledied, ccpvirus, ccpviruscoronavirus, chinaliedpeopledied, chinamustexplain, chinamustpay, chinesebioterrorism, kungflu, makechinapay, milkteaalliance, wholiedpeopledied, wuhanhealthorganisation. The ‘ccp’ in these terms signified Chinese Communist Party, the ruling party of China headed by Xi Jinping, the President of People’s Republic of China. Some of these words also showed the anger towards World Health Organization and termed it as Wuhan Health Organization. This trend shows that China and WHO both are being held responsible for the spread of COVID-19.

Prominent words like virus, trump, pandemic, government, outbreak, pay, communist, propaganda, blame, killed, shame, blame, killing, shit, hell, stupid, lying, lies, die etc. were used to reflect the negative sentiments in the tweets of the people. While words like right, like, good, money, accountable, humanity, responsible, work, organization, great, better, well, global, please, thanks etc. were used to indicate the positive sentiments of the tweets posted by people globally. During the analysis, the terms which are categorized as profanity terms were also analyzed and frequent usage of the profane words in the tweets was observed. The list of these fuck, shit, hell, fucking, ass, crap, screw, fucks, bastards, bitch, bastard, Nazis, ahole and nazi. These words reflected the disgust

emotions of the tweeples.

Overall, the results obtained during the analysis of the tweets, it can be clearly stated that the sentiments of the people tweeting about so called Chinese Virus has been mostly negative. The use of negative words, combined with a good dosage of profane terms, reflect the emotions of tweeples which is mainly concentrated towards the sense of fear, sadness, anger and disgust. The results indicate the signs of discrimination and racism in the COVID-19 era which has been earlier indicated by Coates [26]. The results obtained in this study further strengthen this fact that there has been a vast increase in the online racism due to COVID-19.

Conclusion and Future Works

In this research work, the analysis of the tweets has been done for evaluation of the cyber racism that has been encountered during COVID-19. For this purpose, tweets were collected which mentioned ChineseVirus, WuhanVirus or ChineseVirusCorona. This work explored that the sentiments of majority of the tweets were negative. Further analysis of emotions associated with the tweets also revealed that there was a sense of fear, anger and disgust amongst the people. Apart from this analysis, there were also slangs which reflected negative sentiments towards China, Wuhan and World Health Organization.

During the analysis, majority of the words used in the tweets were negative which included death, pay, communist, ccp, racist etc. The study also revealed a lot of profane words which justified that the cyber racism has been increasing since COVID-19. For the future works, this work can be extended to analyze the trend of the cyber racism in the coming days.

References

[1]. Steiner P. On the Internet, nobody knows you're a dog. The New Yorker, 1993, 69(20): 61-61 - Open Access Library. Oalib.com. http://www.oalib.com/references/16903023. Published 2020. Accessed May 3, 2020.
[2]. Nakamura L. Digitizing race: visual cultures of the Internet. *Choice Reviews Online*. 2008;45(12):45-6837-45-6837. doi:10.5860/choice.45-6837
[3]. Brown C. WWW.HATE.COM: White Supremacist Discourse on the Internet and the Construction of Whiteness Ideology. *Howard Journal of Communications*. 2009;20(2):189-208. doi:10.1080/10646170902869544
[4]. Clark D, Spanierman L, Reed T, Soble J, Cabana S. Documenting Weblog expressions of racial microaggressions that target American Indians. *J Divers High Educ*. 2011;4(1):39-50. doi:10.1037/a0021762
[5]. Huang C, Wang Y, Li X et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet*. 2020;395(10223):497-506. doi:10.1016/s0140-6736(20)30183-5
[6]. Trump D. Twitter. Twitter.com. https://twitter.com/realDonaldTrump/status/1239685852093169664. Published 2020. Accessed May 3, 2020.
[7]. Spit On, Yelled At, Attacked: Chinese-Americans Fear for Their Safety. Nytimes.com. https://www.nytimes.com/2020/03/23/us/chinese-coronavirus-racist-attacks.html. Published 2020. Accessed May 3, 2020.
[8]. Aratani L. 'Coughing while Asian': living in fear as racism feeds off coronavirus panic. the Guardian. https://www.theguardian.com/world/2020/mar/24/coronavirus-us-asian-americans-racism. Published 2020. Accessed May 3, 2020.
[9]. Trump's new fixation on using a racist name for the coronavirus is dangerous. Vox. https://www.vox.com/2020/3/18/21185478/coronavirus-usa-trump-chinese-virus. Published 2020. Accessed

May 3, 2020.

[10]. Zheng, Y., Goh, E., & Wen, J. The effects of misleading media reports about COVID-19 on Chinese tourists' mental health: a perspective article. *Anatolia*. 2020 :*31*(2): 337-340. Published 2020.

[11]. Zhu, H. Countering COVID-19-related anti-Chinese racism with translanguaged swearing on social media. *Multilingua, 1*(ahead-of-print). Published 2020.

[12]. Rovetta, A., & Bhagavathula, A. S. Covid-19-related web search behaviors and infodemic attitudes in Italy: Infodemiological study. JMIR Public Health Surveill, 2020, 6(2), e19374.

[13]. Budhwani H, Sun R. Referencing the novel coronavirus as the "Chinese virus" or "China virus" on Twitter: COVID-19 stigma. J Med Internet Res. 2020: 22(5). doi:10.2196/19301

[14]. Shirsat V, Jagdale R, Shende K, Deshmukh S, Kawale S. Sentence Level Sentiment Analysis from News Articles and Blogs using Machine Learning Techniques. *International Journal of Computer Sciences and Engineering*. 2019;7(5):1-6. doi:10.26438/ijcse/v7i5.16

[15]. Uma Ramya V, Thirupathi Rao K. Sentiment Analysis of Movie Review using Machine Learning Techniques. *International Journal of Engineering & Technology*. 2018;7(2.7):676. doi:10.14419/ijet.v7i2.7.10921

[16]. M A, Ravikumar P. Survey: Twitter data Analysis using Opinion Mining. *Int J Comput Appl*. 2015;128(5):34-36. doi:10.5120/ijca2015906553

[17]. Maindola P, Singhal N, Dubey A. Sentiment Analysis of Digital Wallets and UPI Systems in India Post Demonetization Using IBM Watson. In: *International Conference On Computer Communication And Informatics (ICCCI)*. Chennai: IEEE; 2020:1-6.

[18]. Wang Y, Taylor J. Coupling sentiment and human mobility in natural disasters: a Twitter-based study of the 2014 South Napa Earthquake. *Natural Hazards*. 2018;92(2):907-925. doi:10.1007/s11069-018-3231-1

[19]. Kim, E. H. J., Jeong, Y. K., Kim, Y., Kang, K. Y., & Song, M. Topic-based content and sentiment analysis of Ebola virus on Twitter and in the news. *Journal of Information Science. 2016: 42*(6): 763-781. doi: 10.1177/0165551515608733

[20]. Mamidi, R., Miller, M., Banerjee, T., Romine, W., & Sheth, A. Identifying key topics bearing negative sentiment on Twitter: insights concerning the 2015-2016 Zika epidemic. *JMIR Public Health and Surveillance*. 2019: *5*(2): e11036. doi: 10.2196/11036.

[21]. Roccetti, M., Marfia, G., Salomoni, P., Prandi, C., Zagari, R. M., Kengni, F. L. G. & Montagnani, M. Attitudes of crohn's disease patients: Infodemiology case study and sentiment analysis of facebook and twitter posts. *JMIR public health and surveillance*. 2019: *3*(3): e51. doi: 10.2196/publichealth.7004

[22]. Rinker T. Calculate Text Polarity Sentiment [R package sentimentr version 2.7.1]. Cran.r-project.org. https://cran.r-project.org/package=sentimentr. Published 2019. Accessed May 3, 2020.

[23]. Deng, H., Wang, Q., Turner, D. P., Sexton, K. E., Burns, S. M., Eikermann, M., & Houle, T. T. Sentiment analysis of real-world migraine tweets for population research. *Cephalalgia Reports*. 2020:*3*. doi: https://doi.org/ 10.1177/ 2515816319898867

[24]. Ouyang, Y., & Waterman, R. W. Trump Tweets: A Text Sentiment Analysis. In *Trump, Twitter, and the American Democracy*. Palgrave Macmillan, Cham. 2020: 89-129.

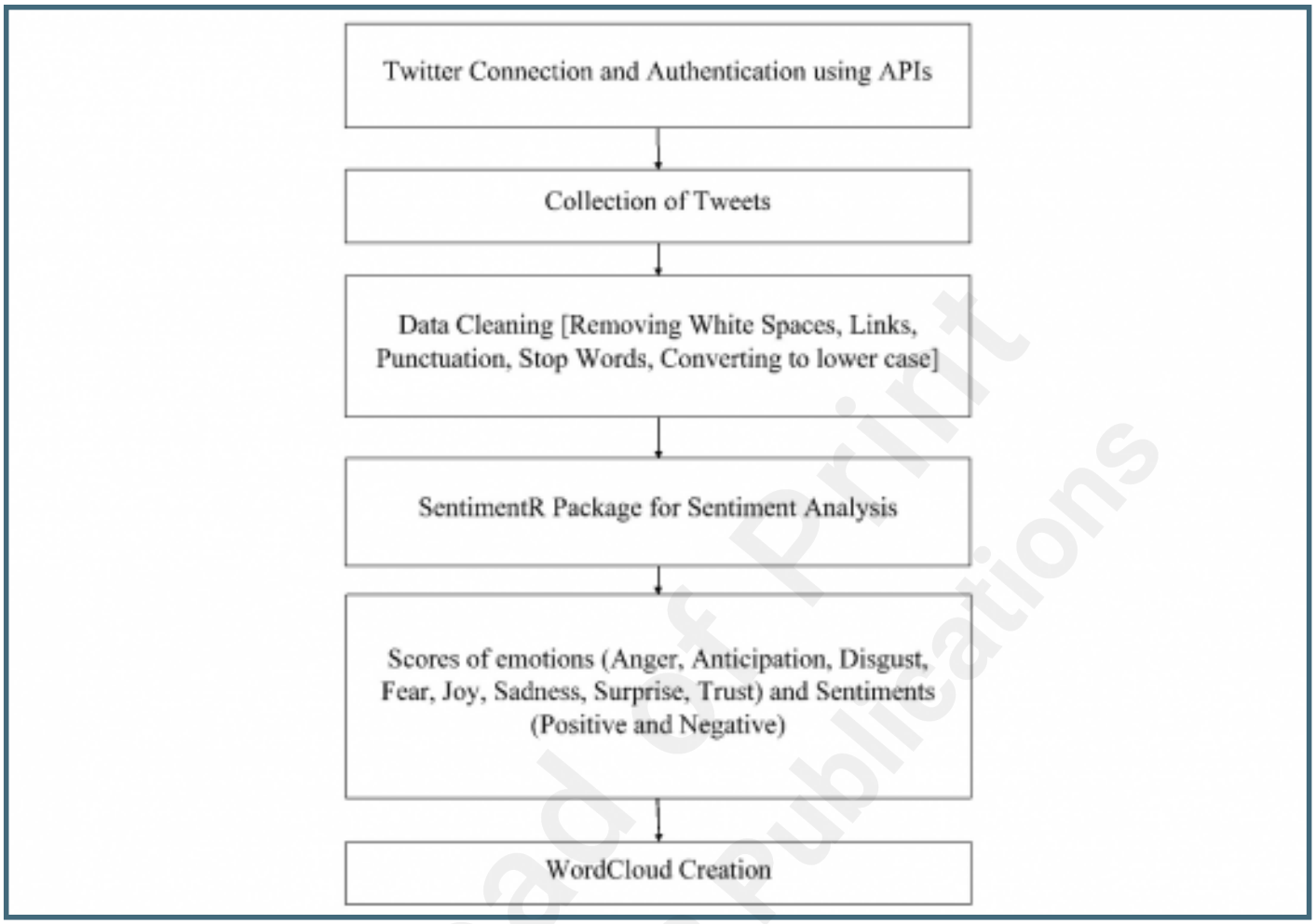
[25]. Naldi, M. A review of sentiment computation methods with R packages. *arXiv preprint. 2019:arXiv:1901.08319*.

[26]. Coates M. Covid-19 and the rise of racism. *BMJ*. 2020:m1384. doi:10.1136/bmj.m1384

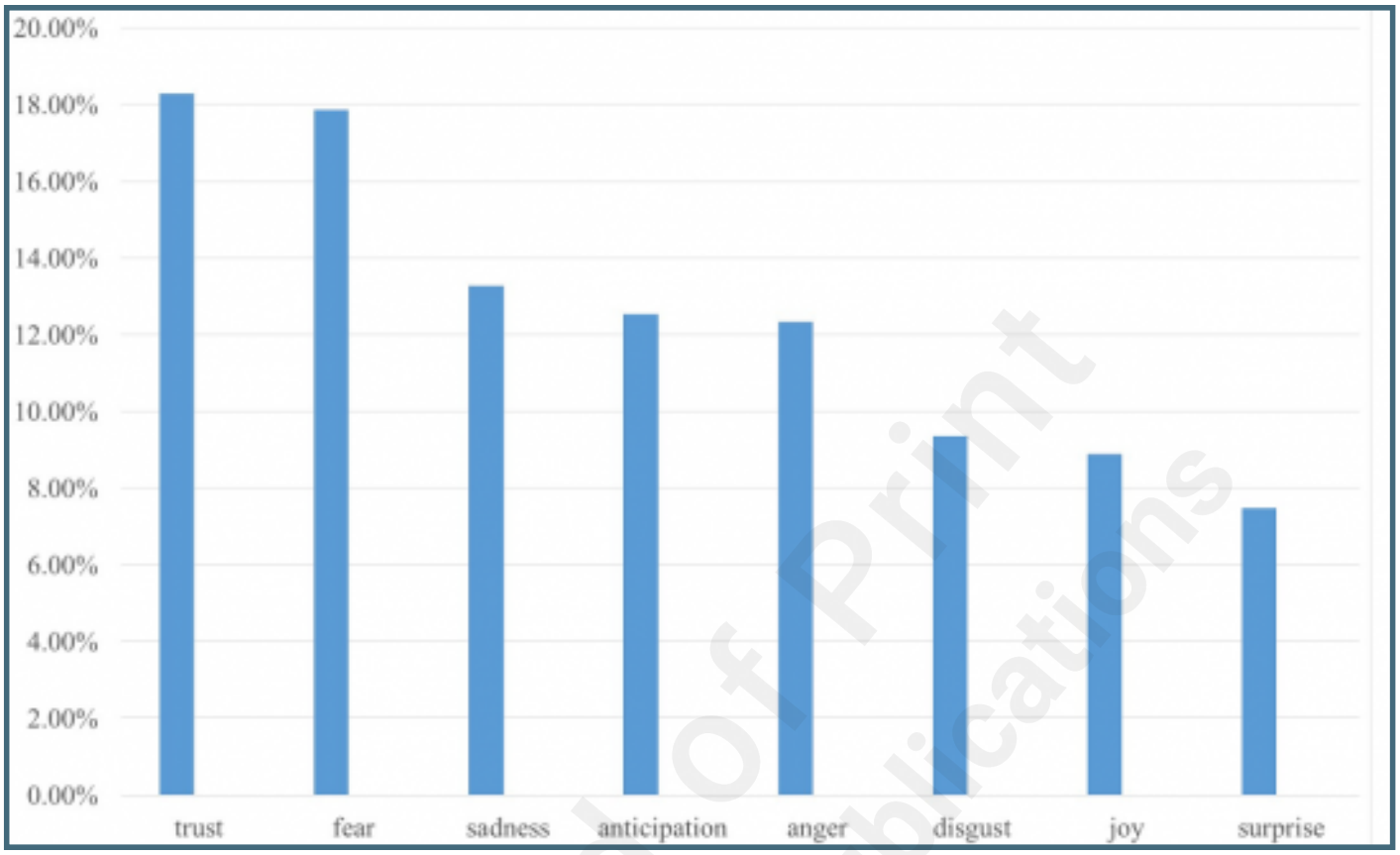
Supplementary Files

Figures

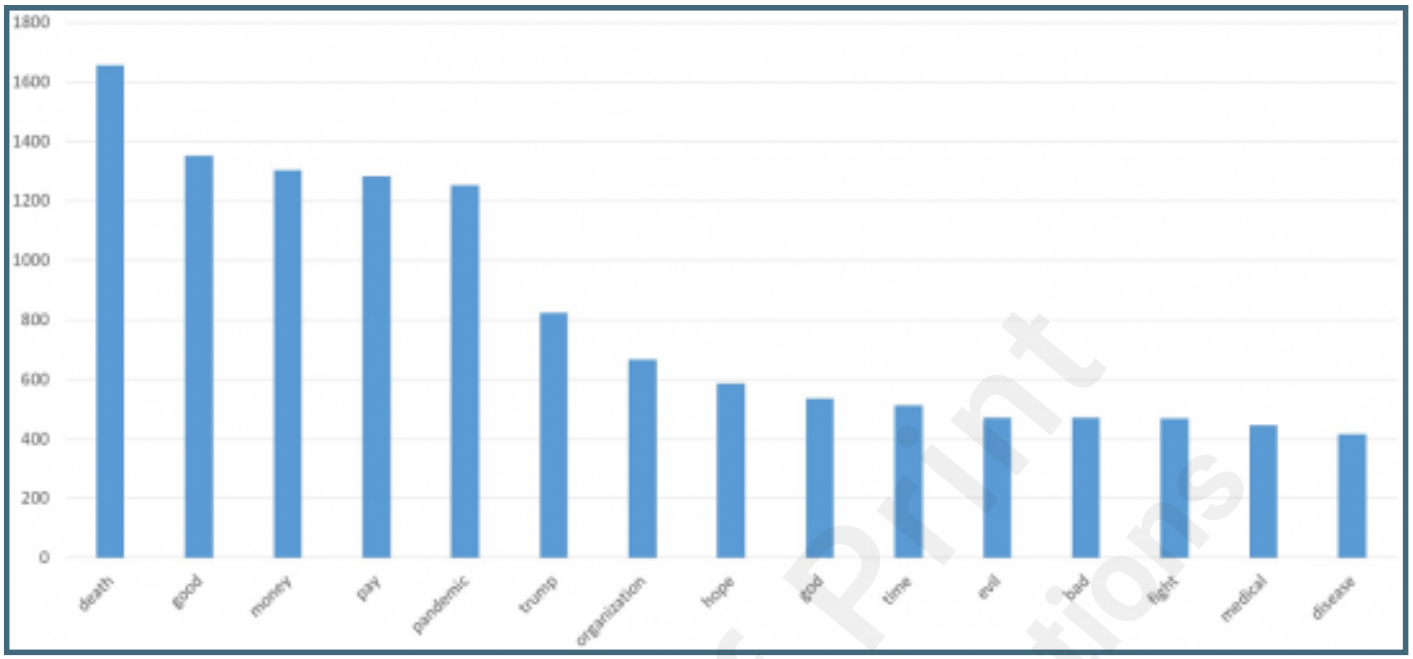
Flowchart for sentiment analysis of tweets.



Emotion analysis of tweets.



Frequency of most used words in the analyzed tweets.



Word cloud of positive and negative sentiment words.