

Analytical model for cultural perception around the world of hate crime in North America during Covid-19 pandemic

Venkata Duvvuri^{1,*}, Gahyoung Lee², Mengyu Yang³ and Yu-Chiao Shaw⁴

^{1,2,3,4} College of Professional Studies, Northeastern University, USA

Abstract.

The world has been hit hard by the coronavirus pandemic that started early in 2020. Hate crime and racism in the US accelerated during the pandemic. The origin of the Covid-19 virus is attributed to China, but without verifiable substance. Additionally, racist incidents like the unfortunate death of George Floyd in 2020 has given rise to Black Lives Matter and other movements. This study focuses on global citizens' attitudes towards US hate crime and racism in social media, namely Twitter, during the pandemic. We compare attitudes from two Asian subcontinents (South and East) and the European continent to North America (US) during this time. Additionally, we study the underlying emotions when people use social media to express views of hate crime and racism in these regions. We study the negativity of the overall sentiment and dig into the dominant emotions. Our analysis uses North America as a baseline, and controls for subjectivity as a moderator. We find that Europeans and South Asians express more personal opinions than factual (5% more subjectivity) compared to North Americans. All regions are more negative in expression than North America, with South Asia being the most negative (16% more polarity). The surprising result is that East Asia is “happier” and “surprised” simultaneously, reflecting mixed political ideologies. Europeans are “angrier” in consonance with the moderating subjectivity factor. Lastly, South Asia is “sad” and “fearful” indicating this region is anxious due to the prevalence of hate crime in the developed nations.

Keywords: Covid-19, hate crime, racism, NLP, subjectivity, polarity, emotions

1. Introduction

Recent research (Piatkowska et. al, 2020) has analyzed this trend in US and attributes such crimes to the threat posed by the subordinate group to the privileged group. Additionally, politics, minority population distribution and economic competition exacerbate these issues (Disha et al. 2011) (Green et al., 1998) (Lyons, 2008). American history is ripe with racism for over a hundred years (Harris, D. A., 2006). With technological advancement one can aim to eliminate hate crime like never before. While minority communities have been targeted as the victims of hate speech in North America, the US government should lead the way to eliminate hate crime in every quarter of the world.

4th WORLD CONFERENCE ON SOCIAL SCIENCES & HUMANITIES

18-20 March 2022 Prague, Czech Republic



The world has been hit by a great pandemic caused by the coronavirus. There is reason (CDC, 2020) to believe that this virus spread from animals to humans in Wuhan's open-air "wet markets" that sell fish and raw meat to consumers. While this is yet to be proved, the net result is a series of crimes against the Asian minority (Gover et al., 2020), (Chelsea & Hansen, 2021). However, the hate crime and racism have spread to non-Asians, including Blacks and Hispanics (Wenger, M. R., & Lantz, B., 2021). This motivates this study of its global perceptions. Specifically, we focus on perceptions outside the North America during the pandemic. With technological innovation, one can now use advanced analytical methods to evaluate perceptions of hate crime and racism. Racist hate crime is a complex issue, but the importance of evaluating its perception in other regions is desirable for several governments. Many choose Twitter to express opinion freely, with 90% of its profiles open to the public (Nguyen et al., 2021). Few studies focus on hate crime perception alone (Woo et al, 2021), but none studied local hate crime perceptions outside the North America (US) using advanced analytical methods.

This paper is organized as follows: Section 1 provides the context of the problem; Section 2 presents a literature survey; Section 3 lays out the methodology of the study; Section 4 analyzes data validity of public speech in racism and hate crime; Section 5 presents the experimental measurements and results, and finally we conclude in Section 6.

2. Prior Work

We evaluate recent papers in the literature since the 2020s (the inception of Covid-19 pandemic), and in Table 1 we compare these studies to our own.

(Croucher et al., 2020) study the relationship between anti-Asian sentiment prevailing in social media during Covid-19 times and observe the acceleration of hate crime during the pandemic.

(Mueller and Schwarz, 2021) investigate the influence of social media on the refugee crisis and resulting anti-refugee sentiment in Germany.

(Williams, 2020) studies the relationship between online hate crime and offline instances.

(Gover et al., 2020) study how hate crimes against Asians during Covid-19 times have propagated to "other" Asian Americans.

(Nguyen et al., 2021) study and evaluate the negative sentiment against Asians in social media.

(Thelwall et al, 2021) study the killing of George Floyd in the US by inspecting phenomena like Black Lives Matter (BLM). They find that BLM accelerated the negative sentiment and introduced additional concerns in social media.

(Mansoor et al, 2020) study the sentiment across time during the Covid-19 pandemic globally. But their study is limited to temporal characteristics.

None of the studies above focus on the perception of US hate crime and racism outside that country. The differences in perceptions of hate crimes in different regions are significant, and as far we are aware only our study analyzes these differences using quantitative metrics like sentiment, polarity and emotions, controlling for moderating factors like subjectivity.

4th WORLD CONFERENCE ON SOCIAL SCIENCES & HUMANITIES

18-20 March 2022 Prague, Czech Republic



Table 1: Literature Survey

Paper	Topic	Dimension	Method
Croucher et al., 2020	Anti-Asian	Time	Survey
Mueller and Schwarz, 2021	Anti-Refugee	Geo (Germany)	Facebook
Williams, 2020	Anti-Asian	Geo (UK)	Twitter
Gover et al., 2020	Anti-Asian	Geo (US)	Qualitative
Nguyen et al., 2021	Anti-Asian	Geo (US)	Twitter
Thelwall et al., 2021	BLM	Time	Twitter
Mansoor et al., 2020	Global	Time	Twitter
Duvvuri et al., 2022	Anti-Asian, BLM	Geo (non-US)	Twitter

Table 2: Speech Topic Studied

	Hate Crime	Racism
Hashtags	# chinesevirus	# BLM
	# hate crime	# racism
	# asian hate crime	
	# stopasianhate	

Table 3: Data Collection

Region	Number of Records	Collection Period
North America (baseline)	2330	Jan15 – Feb 06
Europe	4861	Jan15 – Feb 06
East Asia	809	Jan15 – Feb 06
South Asia	351	Jan15 – Feb 06

3. Methodology

We collected public opinion tweets globally from Twitter (8350+ tweets) on various topics related to hate crime and racism arising during the pandemic, especially in the omicron outbreak of the pandemic (Jan 15, 2022–Feb 06, 2022). We mine several #hashtags listed in Table 2 and verify the data validity by looking at word cloud patterns and statistical plots (Figures 2). This helped us to understand the correlation between hate speech in the public domain and valid word usage patterns in Section 4. We leverage public Twitter APIs by configuring weekly geographic and topic variables in Twitter.

Our study is limited to common expressions of racist and hateful opinions and defers the localized and language-specific expressions to future study. Additionally, we do not account for obfuscation methods to avoid free speech detection, namely cross-language characters or irony in expressions. Finally, we filter out all the re-tweets to avoid amplifying the studied sentiment.

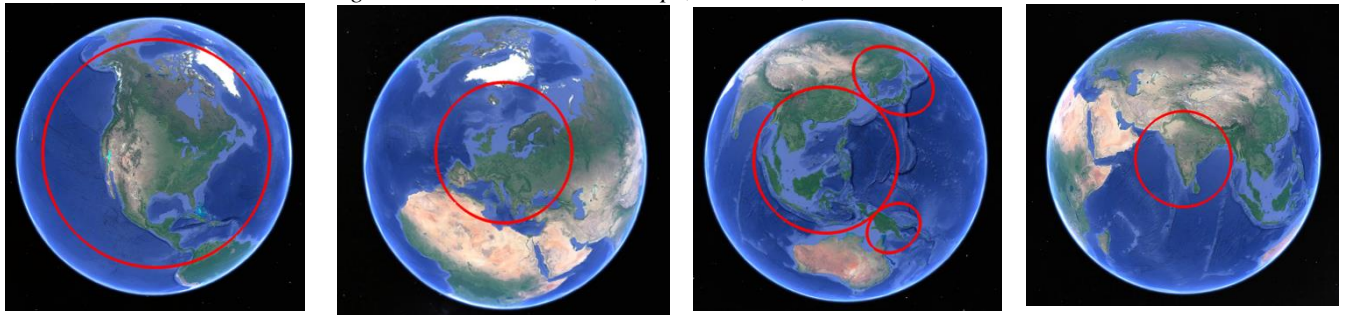
4th WORLD CONFERENCE ON SOCIAL SCIENCES & HUMANITIES

18-20 March 2022 Prague, Czech Republic



In implementation, we leverage open-source Natural Language processing (NLP) packages like Natural Language Toolkit (nltk) and spaCy to detect the patterns relating to the racist posts in Twitter. Nltk package provides a large language dictionary database with different lexical terms, mapping them with the machine learning models. We centrally collect and aggregate weekly #hashtag data in MongoDB to process structured and unstructured data. Additionally, we use spaCy, an open-source software library for advanced natural language preprocessing. Finally, we use SentimentIntensityAnalyzer, TextBlob and text2emotion open-source packages to apply statistical natural language processing to quantify hate crime with NLP-based expressivity metrics: sentiment, polarity and emotion scores.

Figure 1: North America, Europe, East Asia, South Asia



Source: Google Map

Figure 3: Distribution of Speech by Region

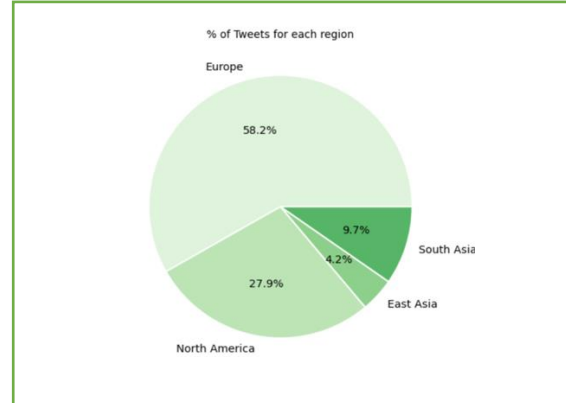
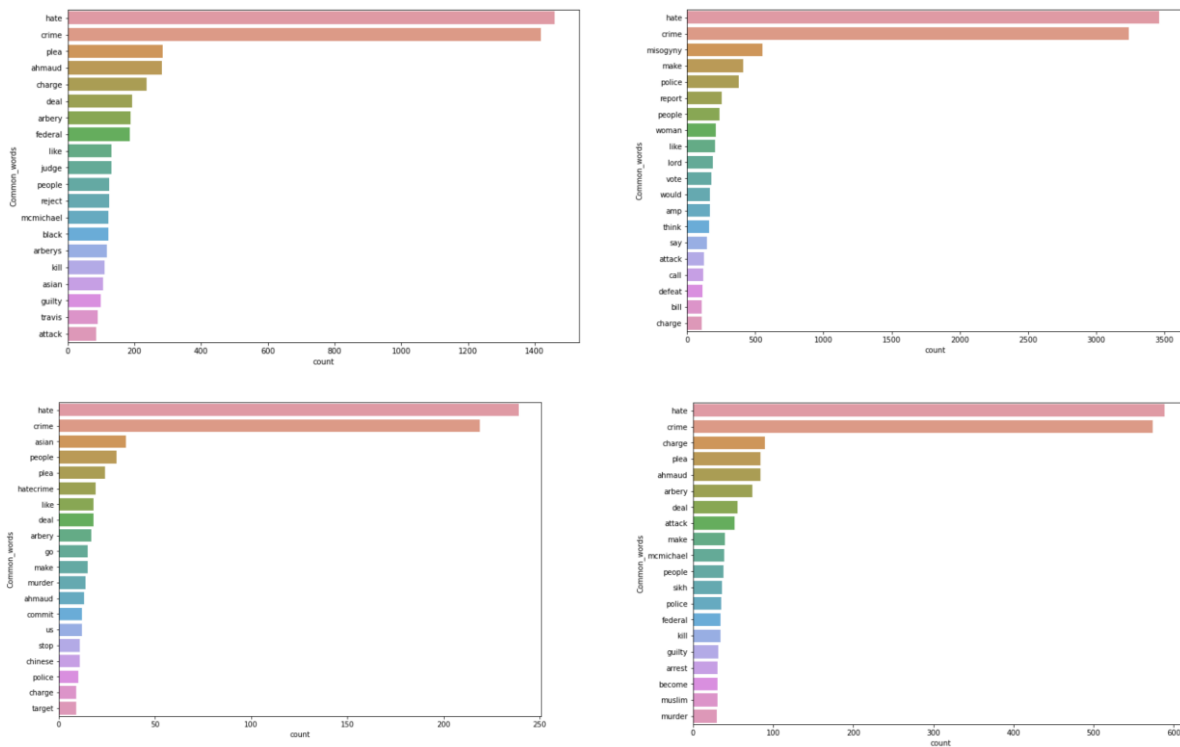


Figure 4: Top words in Speech by Region



5. Experimentation

5.1 Setup

We selected four regions, namely South Asia, East Asia, North America and Europe, by geofencing tweets in Twitter APIs. India, Europe and Asia are the experimental group and North America is the control group. The data was collected once per week, for a total of three weeks during the peak omicron phase of the Covid-19 pandemic. The first collection is from

4th WORLD CONFERENCE ON SOCIAL SCIENCES & HUMANITIES

18-20 March 2022 Prague, Czech Republic



Jan/15/2022 to Jan/22/2022; the second collection is from Jan/23/2022 to Jan/29/2022; the third collection is from Jan/30/2022 to Feb/06/2022. Table 3 reflects cumulative counts for each region which are: 809 for South Asia, 351 for East Asia, 4861 for Europe, and 2330 for North America. The size of the sample gives us several statistically significant results in the metrics.

5.2 Measurement

We use three NLP expressivity metrics for our evaluation of hate speech: sentiment, polarity and emotion scores. We control for the subjectivity factor score qualitatively.

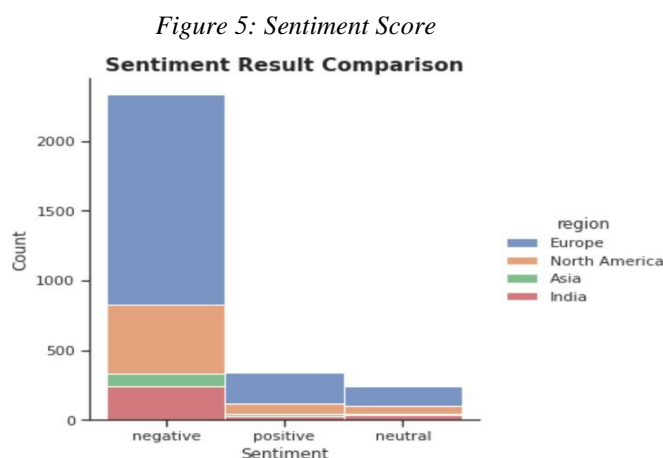
Sentiment score captures the favorability in conversation as positive or negative or neutral (Nasukawa et al., 2003). It is either positive (favorable) or negative (unfavorable) or a neutral opinion towards a text.

Polarity score quantifies the strength of the speech (Nigam & Hurst, 2006). It lies between (-1.0, 1.0). -1.0 defines a negative sentiment and 1.0 defines a positive sentiment, where we defined negative (unpleasant feeling), neutral (feeling neither unpleasant nor pleasant), and positive (pleasant feeling).

Subjectivity score quantifies the amount of personal opinion vs factual information contained in the text (Wiebe, 2000). Subjectivity lies between (0.0, 1.0). Higher subjectivity means that the text contains personal opinions rather than factual information.

We use SentimentIntensityAnalyzer package to capture sentiment, polarity and subjectivity. We leverage text2emotion package to evaluate latent human emotions behind speech. The package classifies text tone into five basic human emotions: Happy, Angry, Surprise, Fear, and Sad.

Our methodology includes detecting emotion from every word in pre-processed text, finding the appropriate words that express emotional sentiment, polarity or feelings, and checking lexical category of each word. Our methods¹ are probabilistic with a higher score indicating more likelihood of message belonging to that emotion.



<https://github.com/merlinmy/Analytical-analysis-of-US-hate-crime-globally>

Figure 6: Polarity Score

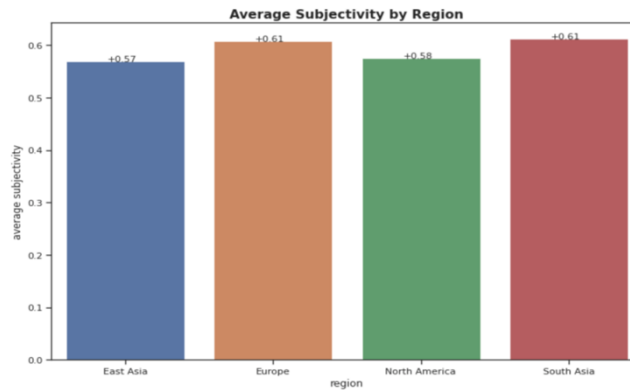
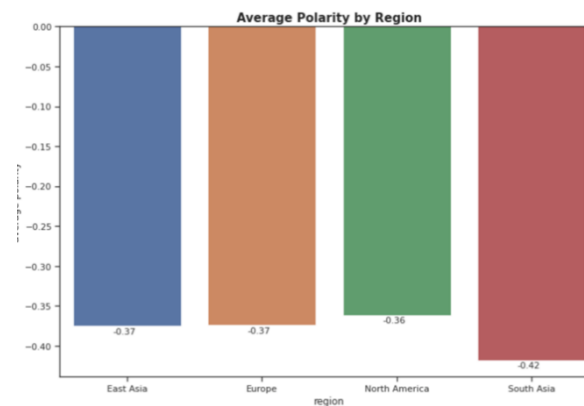


Figure 7: Subjectivity Score



5.3 Results

At a broad level, polarity differences in North America are statistically significant* from the rest of the regions (> 90% confidence level), except in East Asia (>70% confidence level) as per Table 4. Secondly, all subjectivity differences are also statistically significant compared to North America (> 90% confidence level), except in East Asia (> 60% confidence level).

We observe that Europeans & South Asians express more personal opinions than factual (+5.6+%) when compared to North Americans as per Table 4. This accelerates the expression of stronger emotions like “anger” according to our study. Additionally, we find that most regions are more negative in sentiment than North America as per Figure 8. Table 4 reflects that South Asia is the most negative (+15.5%). In Table 5, Europeans are “angrier” in consonance with the hypothesis that democratic countries hope to minimize such hate crimes considering citizen equity as a common goal (Purser, 2017).

4th WORLD CONFERENCE ON SOCIAL SCIENCES & HUMANITIES

18-20 March 2022 Prague, Czech Republic



Surprisingly, East Asia is “happier” as well as “surprised” in Table 5, reflecting the mixed political ideologies (Henders, 2004). The author covers politics of recognition in Japan, politics of cultural difference in China, Buddhist identity in Taiwan, nationalist conflicts in post-Suharto Indonesia, the Philippines and East Timor, and Malaysian ethnic and religious political discourse. In South Asia “sadness”, and “fear” are the dominant emotions, pointing to anxiety (Chatterjee & Katznelson, 2017) in such democracies over recent issues. We concur with the author that such hate crime and racism elevate anxiety in democracies such as India.

Figure 8: Subjectivity Score

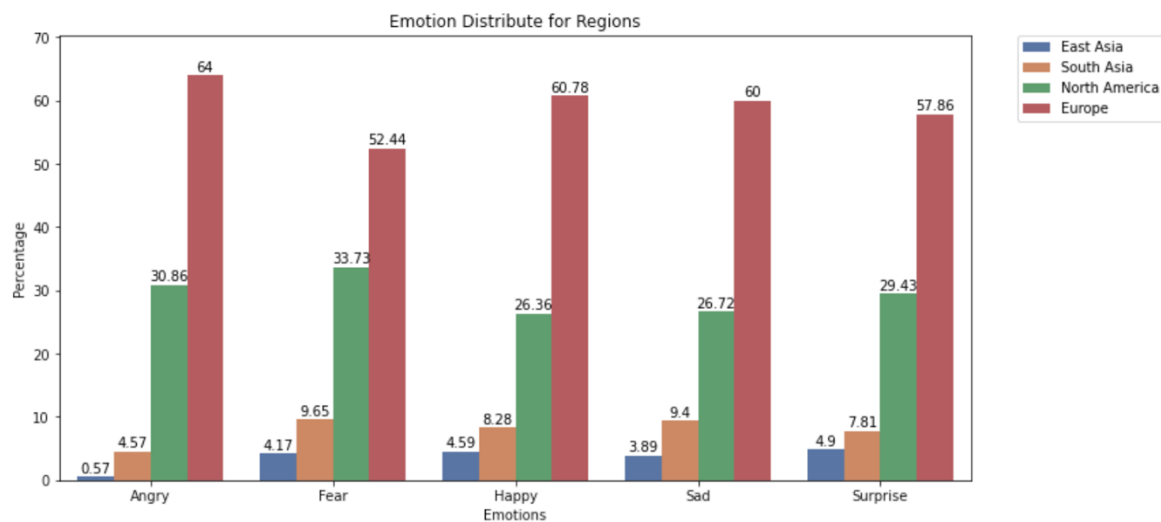


Figure 17 Emotion scores

Table 4: Differences in Polarity and Subjectivity scores

Region	Polarity	Difference	Subjectivity	Difference
North America (baseline)	-0.361	NA	0.575	NA
Europe	-0.373	3.3%*	0.607	5.6%*
East Asia	-0.374	3.6%	0.569	-1.0%
South Asia	-0.417	15.5%*	0.611	6.3%*

* Statistically Significant

Table 5: Differences in Polarity and Subjectivity scores

Region	Angry	Fear	Happy	Sad	Surprise
North America (baseline)	30.86	33.73	26.36	26.72	29.34
Europe	64	52.44	60.78	60	57.86
East Asia	0.57	4.17	4.59	3.89	4.9
South Asia	4.57	9.65	8.28	9.4	7.81

4th WORLD CONFERENCE ON SOCIAL SCIENCES & HUMANITIES

18-20 March 2022 Prague, Czech Republic



6. Conclusion

The Covid-19 pandemic has accelerated hate crime and racism in the North America (US) in the 2020s. The side effects were “anxiety” and “anger” in global geographies on top of common “sadness”. Our study shows that most geographies express negative sentiment over such issues in North America. Europe is more subjective and hence its citizens express amplified emotions like “anger”. South Asian countries are “anxious” over the developments in North America. This could be a major concern for freedom of speech in the free world. Confusingly, East Asians are mixed in sentiment, “happy” and “surprised”. This points to a difference in political ideologies in this region, even within countries. We believe our study can inform foreign policy across the globe. In order to champion freedom in the free world it’s important to factor these differences in perceptions of hate crime and racism.

Acknowledgment

This paper is an output of a collaboration with US Department of State Member - Daniel Lee – and with Global Studies faculty lead in Northeastern University – Fiona Creed.

References

- Croucher. Prejudice toward Asian Americans in the COVID-19 pandemic: the effects of social media use in the United States. *Frontiers in Communication*, 5:39–39, 2020.
- Gover. Anti-Asian hate crime during the COVID-19 pandemic: Exploring the reproduction of in- equality. *American Journal of Criminal Justice*, 45:647– 667, 2020.
- Karsten Mu¨ller and Carlo Schwarz. Fanning the flames of hate: Social media and hate crime. *Journal of the European Economic Association*, 19:2131–2167, 2021.
- L Matthew and Williams. Hate in the machine: Anti-Black and anti-Muslim social media posts as predictors of offline racially and religiously aggravated crime. *The British Journal of Criminology*, 60:20–20.
- Piatkowska, S. J., Messner, S. F., & Hövermann, A. (2020). Black out-group marriages and hate crime rates: A cross-sectional analysis of US metropolitan areas. *Journal of research in crime and delinquency*, 57(1), 105-135.
- Disha, Ilir, James C. Cavendish, and Ryan D. King. 2011. “Historical Events and Spaces of Hate: Hate Crimes against Arabs and Muslims in Post-9/11 America.” *Social Problems* 58(1):21-46.
- Green, Donald P., Dara Z. Strolovitch, and Janelle S. Wong. 1998. “Defended Neighborhoods, Integration, and Racially Motivated Crime.” *American Journal of Sociology* 104(2):372-403.
- Lyons, Christopher J. 2008. “Defending Turf: Racial Demographics and Hate Crime against Blacks and Whites.” *Social Forces* 87(1):357-85.

4th WORLD CONFERENCE ON SOCIAL SCIENCES & HUMANITIES

18-20 March 2022 Prague, Czech Republic



- Centers for Disease Control and Prevention. (2020). Coronavirus disease 2019 (Covid-19) situation summary. Retrieved from <https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/summary.html>.
- Chelsea, G., & Hansen, K. (2021). Did Covid-19 lead to an increase in hate crime against Chinese people in London?. *Journal of Contemporary Criminal Justice*.
- Wenger, M. R., & Lantz, B. (2021). Generalized Hate: Bias Victimization against Non-Asian Racial/Ethnic Minorities during the COVID-19 Pandemic. *Victims & Offenders*, 1-24.
- Nguyen, T. T., Huang, D., Michaels, E. K., Glymour, M. M., Allen, A. M., & Nguyen, Q. C. (2021). Evaluating associations between area-level Twitter-expressed negative racial sentiment, hate crimes, and residents' racial prejudice in the United States. *SSM-population health*, 13, 100750.
- Woo, B., Pitner, R., & Wilson, B. (2021). White College Students' Racial Prejudice and Perceptions of Racial Hate Crime. *Journal of interpersonal violence*, 08862605211062987.
- Harris, D. A. (2006). US experiences with racial and ethnic profiling: History, current issues, and the future. *Critical criminology*, 14(3), 213-239.
- Thelwall, Michael and Thelwall, Saheeda, Twitter during COVID-19: George Floyd Opening a Space to Address Systematic and Institutionalized Racism? (January 12, 2021). Available at SSRN: <https://ssrn.com/abstract=3764867> or <http://dx.doi.org/10.2139/ssrn.3764867>
- Mansoor, M., Gurumurthy, K., & Prasad, V. R. (2020). Global sentiment analysis of COVID-19 tweets over time. arXiv preprint arXiv:2010.14234.
- Nasukawa, T., & Yi, J. (2003, October). Sentiment analysis: Capturing favorability using natural language processing. In *Proceedings of the 2nd international conference on Knowledge capture* (pp. 70-77). Chicago
- Nigam, Kamal, and Matthew Hurst. "Towards a Robust Metric of Polarity." *Computing Attitude and Affect in Text: Theory and Applications*. Dordrecht: Springer Netherlands. 265-279. Web. 2006
- Wiebe, J. (2000). Learning subjective adjectives from corpora. *Aai/iaai*, 20(0), 0.
- Purser, E. (2017). *The Criminalisation of Hate Crimes in a Democratic Society: Attempting to Achieve Equality*. *NEL Rev.*, 5, 29.
- Chatterjee, P., & Katznelson, I. (Eds.). (2012). *Anxieties of democracy: Tocquevillean reflections on India and the United States*. Oxford University Press.

**4th WORLD CONFERENCE ON
SOCIAL SCIENCES & HUMANITIES**
18-20 March 2022 Prague, Czech Republic



Henders, Susan J. *Democratization and Identity: Regimes and Ethnicity in East and Southeast Asia*. Lanham, Md: Lexington Books, 2004. Print.