

Motivation

- The Human Rights Watch reported an increase in the harassment of Asians since the start of the COVID-19 pandemic.¹ This rise may have been exacerbated by anti-Asian messages conveyed by prominent world leaders (e.g., former President Donald Trump, former Minister of Education of Brazil Abraham Weintraub).^{2,3}
- The goals of this research are: (1) to identify patterns of anti-Asian racism on Twitter during COVID-19, (2) build a better understanding of racially-motivated online harassment, and (3) shed light on data patterns applicable to predictive models for identifying instances of racially-motivated online harassment.

Methodology

- **Data Collection** - Twitter API is used and scripted with Python to retrieve tweets that include keywords/hashtags such as: "Chinavirus," "MakeChinaPay," "WashTheHate," "Kungflu," and "IAmNotAVirus." Data has been collected from May 26th, 2020 onwards with plans to collect data leading up to the pandemic. Data will be uploaded to MongoDB/Pandas for cleaning, integration, and analysis.
- **Data Analysis** - Aggregation queries will identify properties of keywords such as the number of instances per day, which keywords are most common, and the geographic location (e.g., U.S. state) of anti-Asian tweets. An algorithm developed by Kleinberg⁴ will be used to observe temporal patterns of racial harassment. The data will also be explored in relation to the timeline of notable COVID-19-related events.
- **Data Visualization** - Additionally, data visualizations will be generated to help identify and understand connections and insights emerging from the data.

Preliminary Analysis

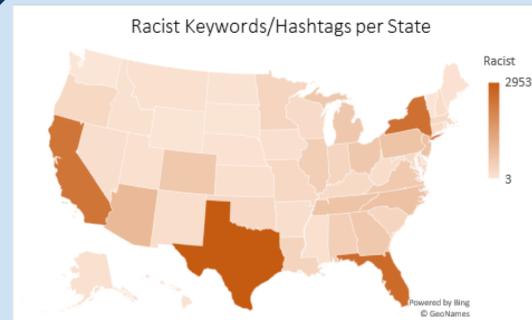


Figure 1. Frequency of racist keywords/hashtags per state. The top 4 states with the highest number of racist keywords/hashtags were Texas (2,953), Florida (2,588), New York (2,480), California (2,385). The 5th highest was Arizona (945).

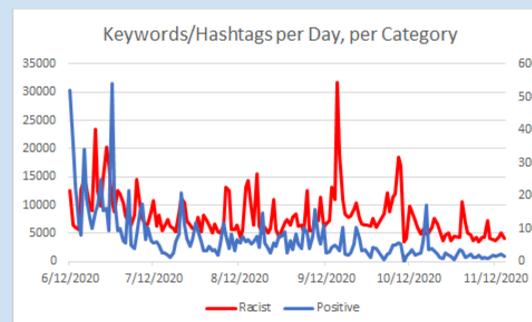


Figure 2. Frequency of negative and positive keywords/hashtags. Keywords/hashtags were categorized as either positive (#WashTheHate) or negative (#Kungflu) and then graphed over the course of several months.

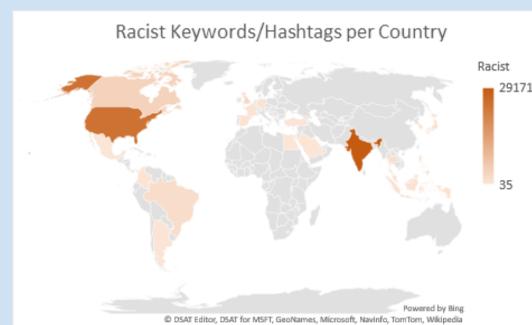


Figure 3. Global view of racist keywords/hashtags. The top 2 countries with the highest number of racist keywords/hashtags were India (29,171) and the United States (24,115). The 3rd highest was Canada (3,988).

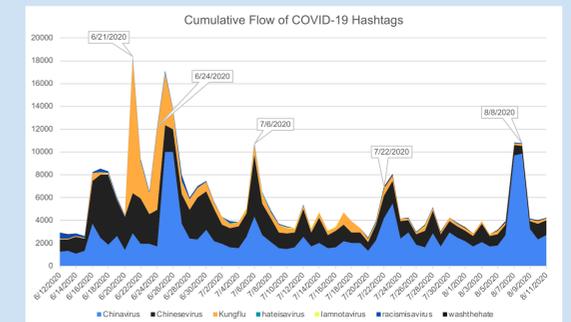


Figure 4. Cumulative flow of COVID-19 keyword/hashtags. The top 3 most common keywords/hashtags were Chinavirus (48.69%), Chinesevirus (34.51%), and Kungflu (14.81%). The 4th highest was hateisavirus (0.11%).

Important Events/Dates

June 20th – Trump’s rally in Tulsa Calls Covid-19 Coronavirus The ‘Kung Flu’

June 23rd – Trump Once Again Calls Covid 19 Coronavirus The ‘King Flu’

July 7th – The number of coronavirus cases in the U.S. surpassed three million. The United States officially notified the United Nations of its withdrawal from the World Health Organization.

July 22nd – Trump Touts Minority Unemployment ‘Before the China Plague’ to Rebut Biden’s Claim He’s ‘First’ Racist President

August 8th – Trump signs orders aimed at extending some pandemic relief after Congress fails to reach a deal. Trump urges Americans to stop politicizing the coronavirus, blames China.

References

- [1] Human Rights Watch (2020, May 12). *Covid-19 Fueling Anti-Asian Racism and Xenophobia Worldwide*. Retrieved from <https://www.hrw.org/news/2020/05/12/covid-19-fueling-anti-asian-racism-and-xenophobia-worldwide>
- [2] U.S. Office of the Press Secretary (Producer). (2020, July 30). *Remarks by President Trump in Press Briefing | July 30, 2020* [Transcript]. Retrieved from www.whitehouse.gov/briefings-statements/remarks-president-trump-press-briefing-july-30-2020/
- [3] France-Press, A. (2020, April 6). *China outraged after Brazil minister suggests Covid-19 is part of 'plan for world domination'*. Retrieved from <https://www.theguardian.com/world/2020/apr/07/china-outraged-after-brazil-minister-suggests-covid-19-is-part-of-plan-for-world-domination>
- [4] Jon Kleinberg. 2003. *Bursty and hierarchical structure in streams*. *Data Mining and Knowledge Discovery* 7, 4 (2003), 373–397. <https://doi.org/10.1023/A:1024940629314>.

