Highlighting Ethnic Biases in COVID-19 Articles

Kanika Chopra

December 5, 2022

Background

- Language is able to capture stereotypes in literature, even subtly
- Word embeddings can be used to quantify historical trends and social change
- Word frequencies, parts of speech and semantic categories can display bias in text

Why COVID?

1. COVID-19 has enabled to spread of racism and enabled xenophobia resulting in Asian Americans being vulnerable to hate crimes
2. Xenophobia and bigotry resulted in the formation of:
   - Asian-American Pacific Islander Equity Alliance (AAPI Equity)
   - Chinese for Affirmative Action (CAA)
   - Stop AAPI Hate coalition in March 2020
3. Many articles describing various hate-crime incidents globally

Plan

Problem

Investigate the COVID-19 pandemic's influence on the perception of Asians in global news articles

1. Quantify the association with neutral word groups to show bias towards Asians
2. Compare the sentiment behind the top 10 most-biased adjectives for Asians (relative to Whites)

Data

Word2Vec Embeddings

https://www.overleaf.com/project/63855eeabdc30841a7d4dc09Aylien COVID-19 News DataJanuary - July 2020Sampled 20, 000 articles per monthGlobal news sources

Word Lists

Group Words

- Asian last names
- White last names

Comparison Words

- COVID Terms
- Hate Crime
- Outsider Adjectives
- General Adjectives

Quantify Embedding Bias

Two methods for comparison:

1. Compute group representational vector for each ethnicity
2. Compute group representational vector for categories

Group Ethnicities

Let A and W be the number of Asian and White last names respectively. Let w_i represent a word in the given group and c_i represent a word from a category.

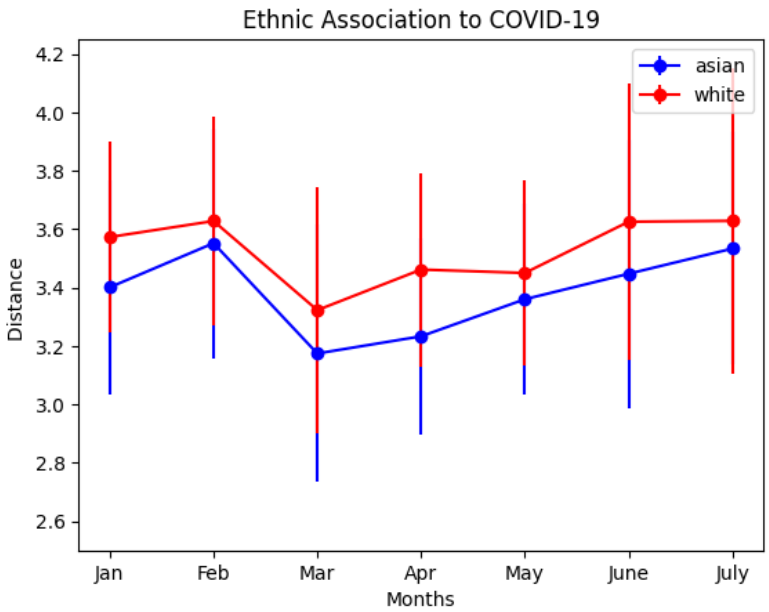
1. $\mu_{\text{asian}} = \frac{1}{A} \sum_{i=1}^{A} w_i$
2. $\mu_{\text{white}} = \frac{1}{W} \sum_{i=1}^{W} w_i$
3. $\delta_{\text{asian},i} = ||\mu_{\text{asian}} - c_i|| \quad \forall i$
4. $\delta_{\text{white},i} = ||\mu_{\text{white}} - c_i|| \quad \forall i$
5. $\text{bias} = \frac{1}{n} \sum_{i}^{n} (\delta_{\text{asian},i} - \delta_{\text{white},i})$

Group Categories

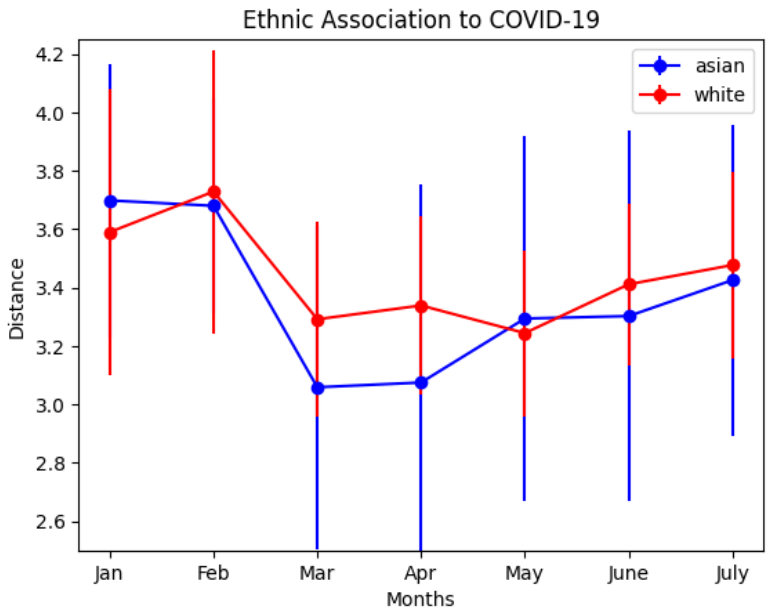
Let A and W be the number of Asian and White last names respectively. Let w_i represent a word in the given group and c_i represent a word from a category.

1. $\mu_{\text{covid}} = \frac{1}{n} \sum_{i=1}^{c} c_i$
2. $\delta_{\text{asian}} = \frac{1}{A} \sum_{i}^{A} ||w_{\text{asian},i} - \mu_{\text{covid}}||$
3. $\delta_{\text{white}} = \frac{1}{W} \sum_{i}^{W} ||w_{\text{white},i} - \mu_{\text{covid}}||$
4. $\text{bias} = \delta_{\text{asian}} - \delta_{\text{white}}$

COVID-19 Terms



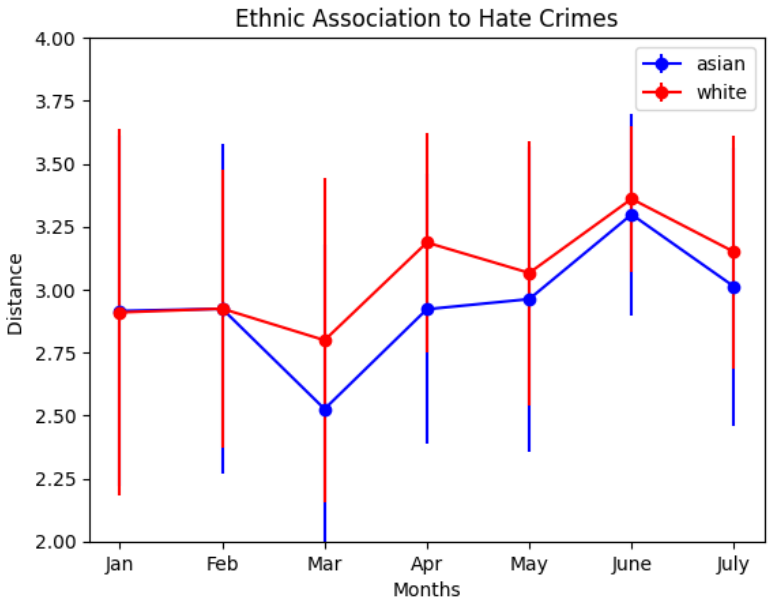
(a) Ethnicity Group Comparison



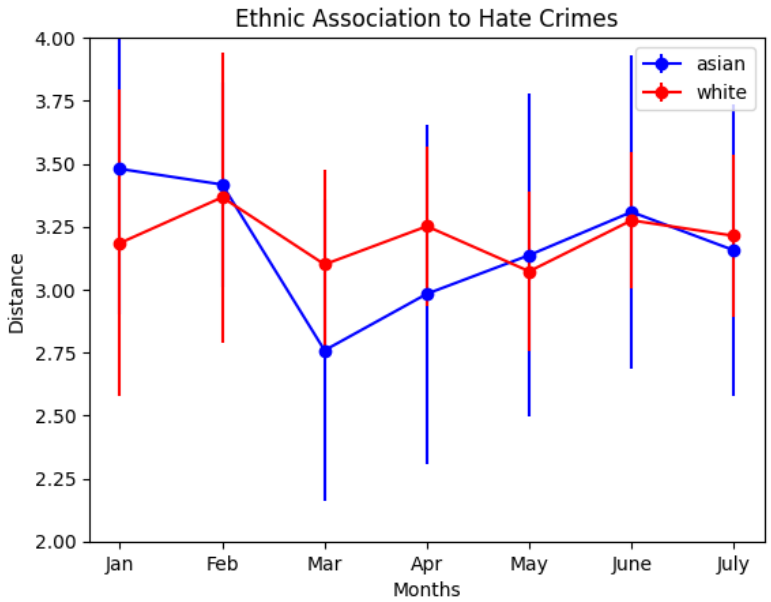
(b) Category Group Comparison

Figure: Association with COVID-19 Terms

Hate Crimes



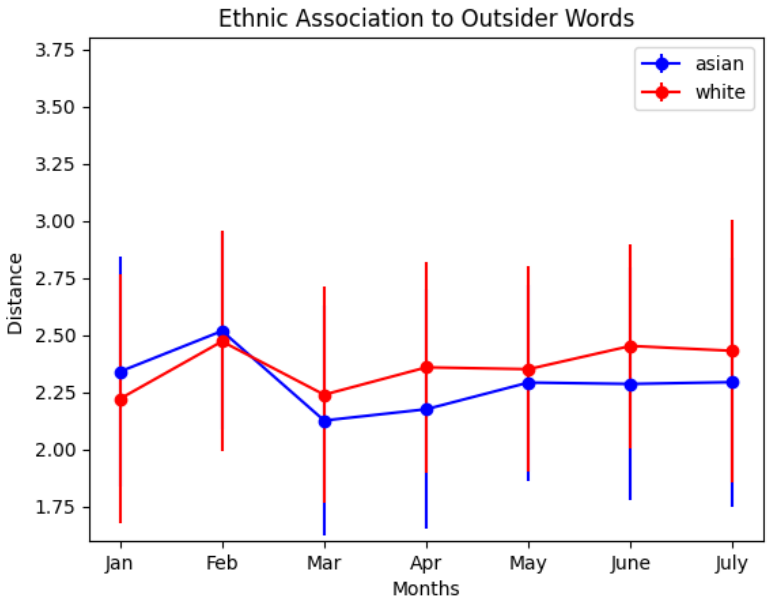
(a) Ethnicity Group Comparison



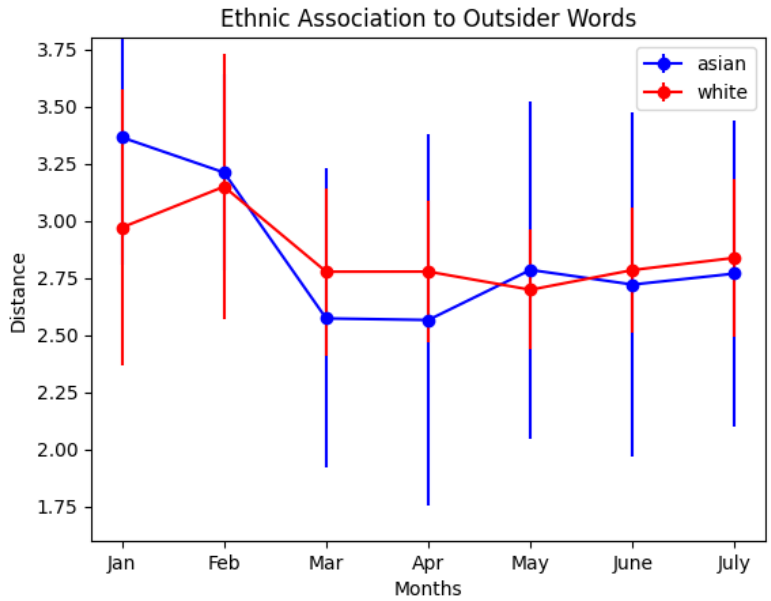
(b) Category Group Comparison

Figure: Association with Hate Crimes

Outsider Adjectives



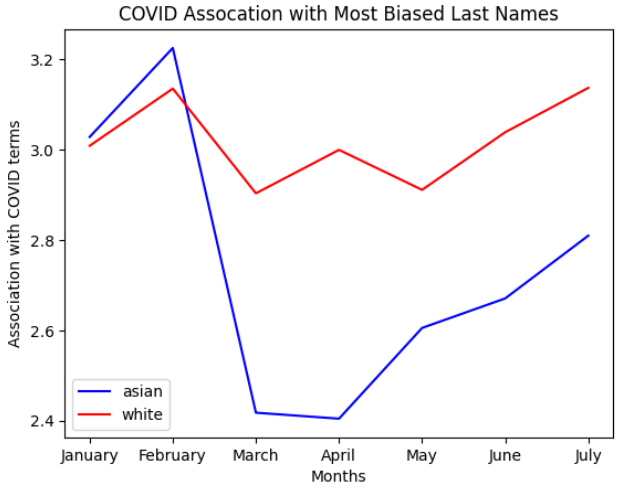
(a) Ethnicity Group Comparison



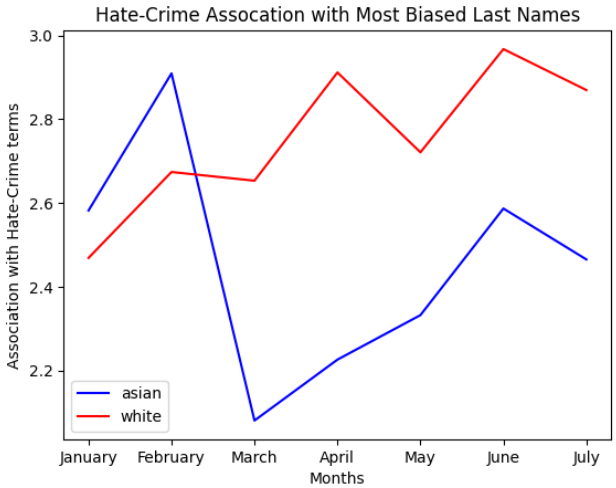
(b) Category Group Comparison

Figure: Association with Outsider Adjectives

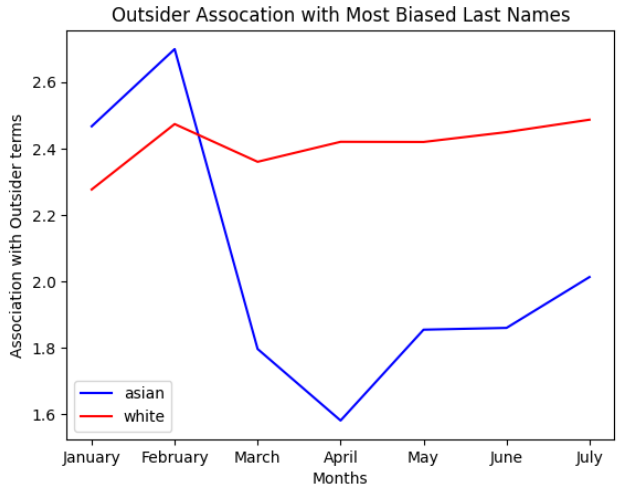
Most Biased Last Names



COVID Terms



Hate Crimes



Outsider Adjectives

Comparing Sentiment Scores

Explore association with 325 adjectives over time

1. Compute average association with μ_{asian} and μ_{white}
2. Compute embedding bias for each adjective
3. Extract top 10 most biased adjectives for each month
4. Compute average compounded sentiment score for Asians and White last names

Top 10 associated words

Jan	Mar	May	July
Protective	Artificial	Insulting	Malicious
Praising	Transparent	Malicious	Cooperative
Cerebral	Informal	Transparent	Assertive
Affected	Sensitive	Cooperative	Arbitrary
Suspicious	Cultured	Praising	Outrageous

Table: Most Biased Adjectives to Asian Ethnicities

Sentiment Analysis

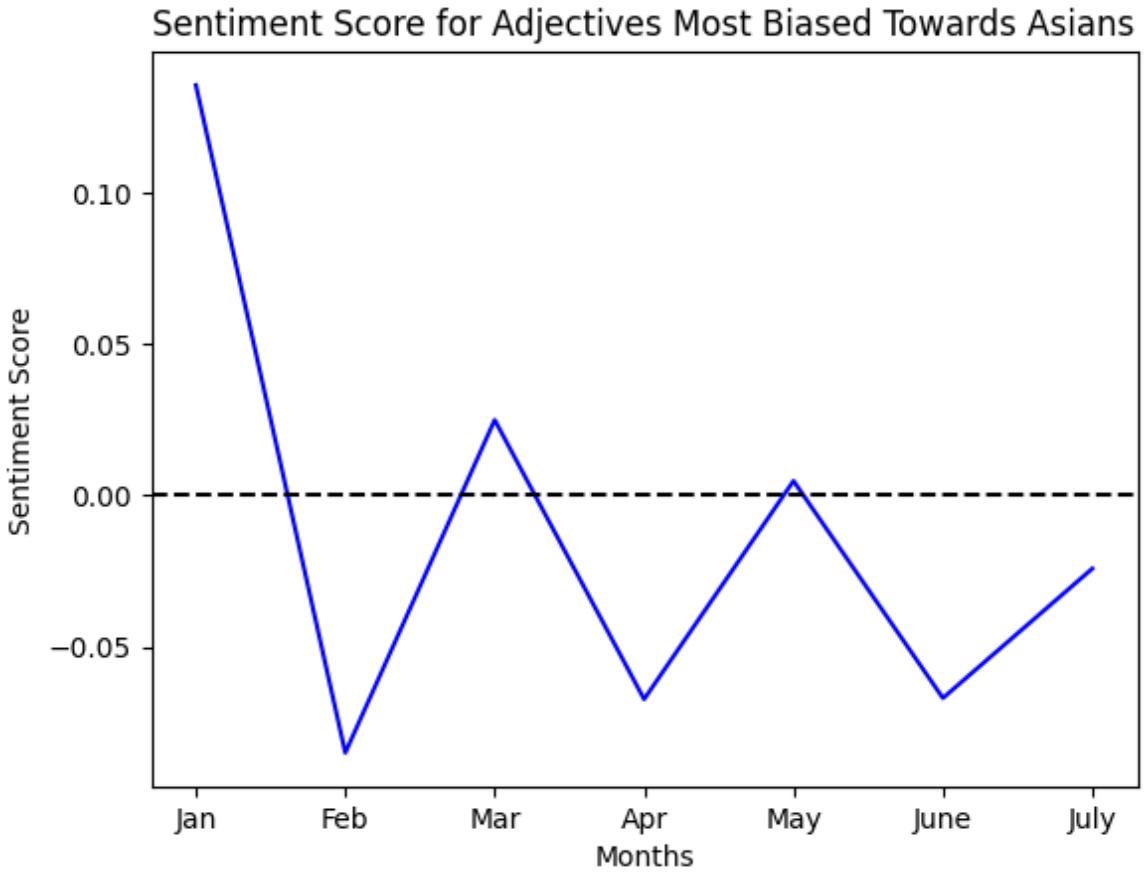


Figure: Sentiment of Most Biased Adjectives

Limitations

- Lack a comparison of the perception before COVID
- Results are dependent on word lists
- No external metrics for validation
- Small data set for statistical tests
- Sentences that include White and Asian last names can cause confusion

Future Work

- Investigate country-wise bias by segmenting news sources
- Compare bias with hate-crime to hate-crime statistics in a country
- Investigate association with blame-related words