Josiah Kim

SODA 496

Professor Ladd

03 May 2021

Donald Trump's Anti-Asian Rhetoric and Anti-Asian Hate Crimes

Overview and Hypothesis

Recent events highlighted by social media suggest that there has been a significant increase in Anti-Asian hate crimes. The Anti-Defamation League's Report on Online Hate and Harassment 2021[1], suggest that "the biggest jump in severe harassment was reported by Asian-Americans, at 17%, compared to 11% reported a year ago."

Theories continue to float around online about the causal mechanism behind the rise in anti-Asian hate crimes. The consensus among social-media users point to the racist rhetoric of Donald Trump for the rise in anti-Asian hate crimes. The Anti-Defamation League saw an "85% increase in anti-Asian sentiment on Twitter" after Donald Trump's contraction of COVID-19. Furthermore, the report suggests that these online sentiments have made their way to the physical world.

A study of the association of "#covid19" versus "#chinesevirus" and anti-Asian sentiments on Twitter also suggest that Donald Trump's rhetoric spark anti-Asian sentiment[2]. The researchers found that only about a fifth of the tweets with "#covid19" had anti-Asian sentiment compared to half the tweets containing "#chinesevirus" that had anti-Asian sentiment.

---

[1] Found at https://www.adl.org/online-hate-2021

[2] Yulin Hswen, Xiang Xu, Anna Hing, Jared B. Hawkins, John S. Brownstein, Gilbert C. Gee, "Association of "#covid19" Versus "#chinesevirus" With Anti-Asian Sentiments on Twitter: March 9–23, 2020", American Journal of Public Health 111, no. 5 (May 1, 2021): pp. 956-964. https://doi.org/10.2105/AJPH.2021.306154

My theory stems off these reports and states as Donald Trump continues to tweet using anti-Asian (negative Asian sentiment), the more anti-Asian hate crimes Asian communities will experience. More specifically, I expect the sentiment of Asian tweets to be an important indicator of the frequency of anti-Asian hate crimes.

Data and Measurement

I draw from two sources of data. The first dataset I employ is the full archive of Donald Trump's tweets[3]. On January 8th, 2021, Twitter banned Donald Trump's twitter account which disabled access to his account even with the Twitter API. However, this archive allowed access to all his past tweets without the use of the Twitter API. This dataset contains 10 variables with each case representing a tweet from Donald Trump (*N = 56,571*).

The second dataset that I employ is the FBI's Uniform Crime Reporting (UCR) program's hate crime database[4]. This dataset contains 28 variables with each case representing a hate crime (*N = 209, 442*) between the years of 1991 and 2019. Both datasets were merged for the purpose of anaysis.

Figure 1 shows the summary statistics of the merged dataset. There are six variables with each case representing a tweet from Donald Trump (*N = 32,786*). My main outcome variable of interest is the number of anti-Asian hate crimes occurred on that day (variable *frequency*). My main independent variable of interest is the sentiment of a tweet (variable *sentiment*) which indicates whether a tweets published by Donald Trump is negative or positive. Another independent variable of interest is whether the tweet is about Asians (variable *asian*). If the tweet is not about Asians (= FALSE) and if the tweet is about Asians (= TRUE).

[3] Found at https://www.thetrumparchive.com/
[4] Found at https://crime-data-explorer.fr.cloud.gov/pages/home

```
     date              text           textstripped         sentiment          frequency
 Min.   :2015-06-16   Length:32786       Length:32786       Min.   :-1.50570   Min.   :0.0000
 1st Qu.:2017-01-26   Class :character   Class :character   1st Qu.:-0.12127   1st Qu.:0.0000
 Median :2019-06-13   Mode  :character   Mode  :character   Median : 0.05669   Median :0.0000
 Mean   :2018-10-11                                        Mean   : 0.06985   Mean   :0.2603
 3rd Qu.:2020-04-13                                        3rd Qu.: 0.27578   3rd Qu.:0.0000
 Max.   :2021-01-08                                        Max.   : 1.81122   Max.   :4.0000

  asian
 Mode :logical
 FALSE:32065
 TRUE :721
```

Figure 1. Descriptive Statistics of Merged Dataset

Analysis and Findings

Figure 2 displays the top 100 hate crimes reported to the FBI by bias since the inception of the UCR program. We can see that anti-Asian hate crimes is the eighth most occurring hate crime. Compared to anti-Black hate crimes at first, anti-Asian hate crimes occur about one-tenth less.

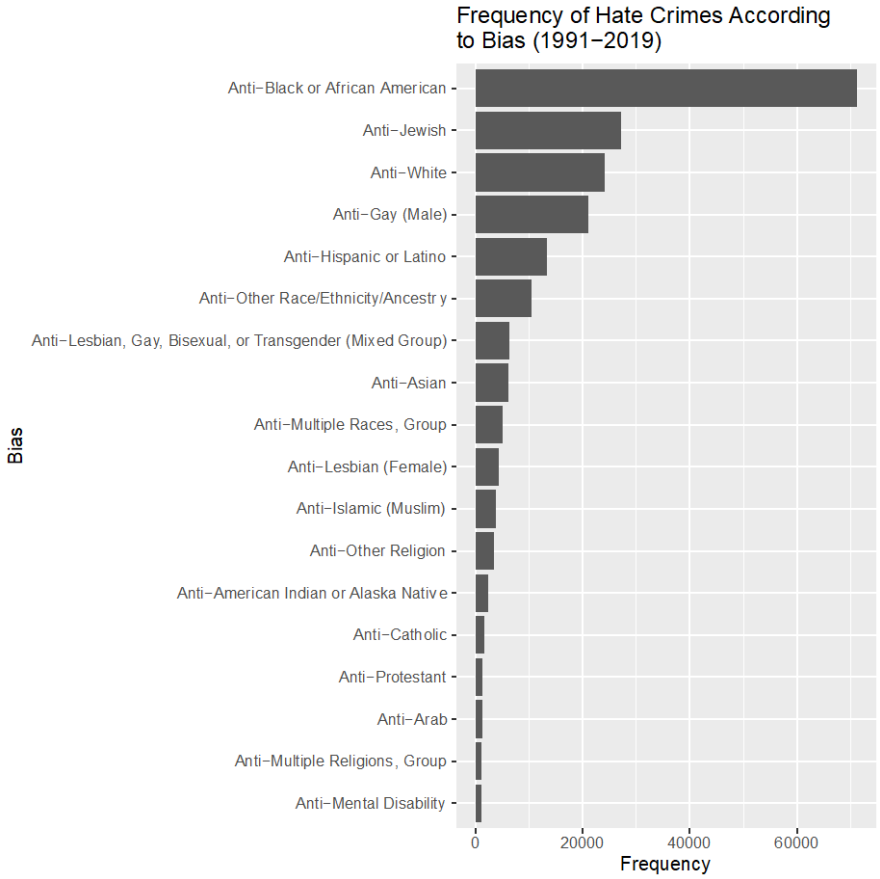


Figure 2. Frequency of Hate Crimes According to Bias (1991-2019)

By decreasing the range of time, figure 2 shows the most occurring hate crimes according to bias starting at the point in time when Donald Trump's announcing his candidacy for president of the United States of America (June 16, 2015). The graph shows that anti-Asian hate crimes are the eleventh most occurring hate crime. In comparison to figure 1, it is suggested that anti-Asian crimes are less occurring than other biases.

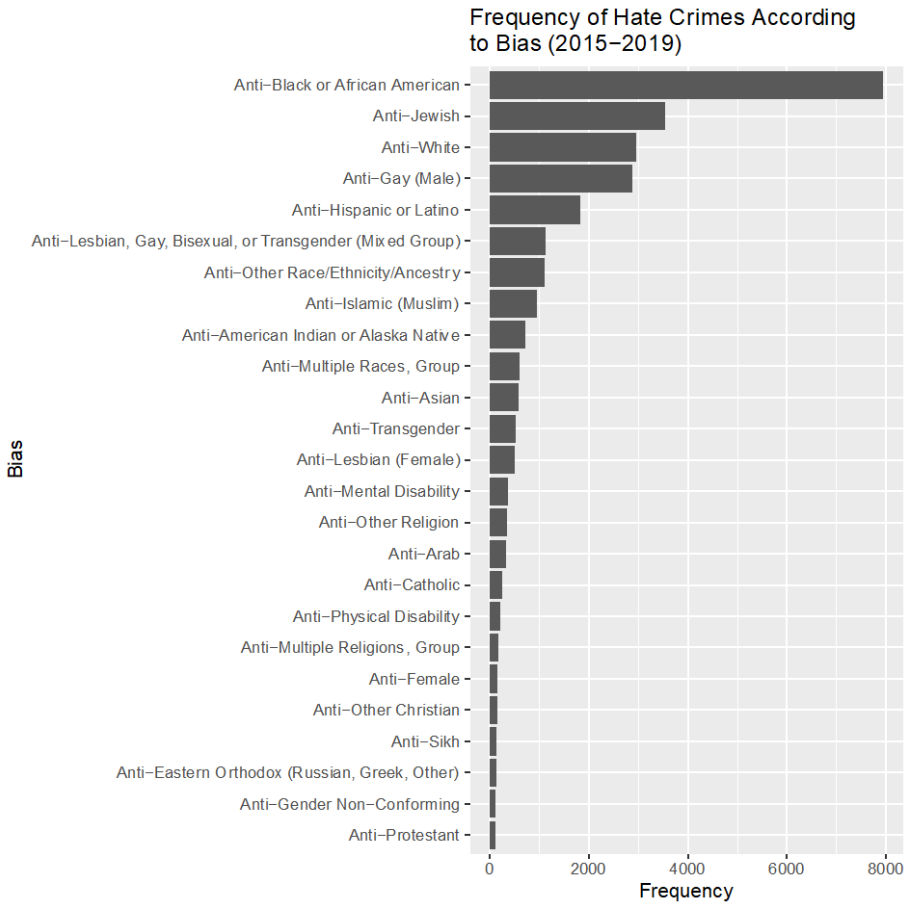


Figure 3. Frequency of Hate Crimes According to Bias (2015-2019)

Figure 4 illustrates a more detailed perspective of anti-Asian hate crimes by showing the occurrences of anti-Asian hate crimes by year since the inception of the UCR program. The greatest number of reported anti-Asian hate crimes occur between the years of 1995 and 1996.

Since 1996, figure 4 depicts a downward trend showing that suggesting that the occurences of anti-Asian hate crimes are decreasing. The portion of the graph in red highlights the period of time starting at the point at which Donald Trump announced his candidacy for presidency. Coincidentally, we can see the first upward trend in anti-Asian hate crimes during this time period as well.

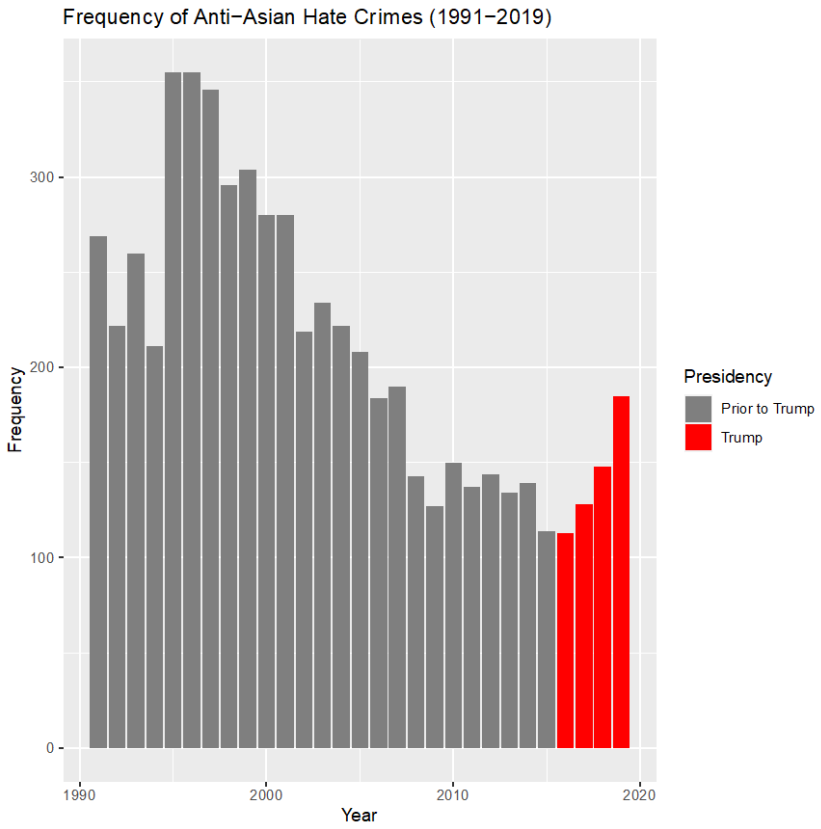


Figure 4. Frequency of Anti-Asian Hate Crimes (1991-2019)

Figure 5 shows the occurrences each year in which Donald Trump tweeted. His first tweet occurred in 2009 with his last tweet occurring in 2021. The red portion highlights the same time period as depicted in figure 4. We see a similar upward trend in tweets during his this time period which may or may not indicate an upward trend of negative sentiment tweets about Asians.

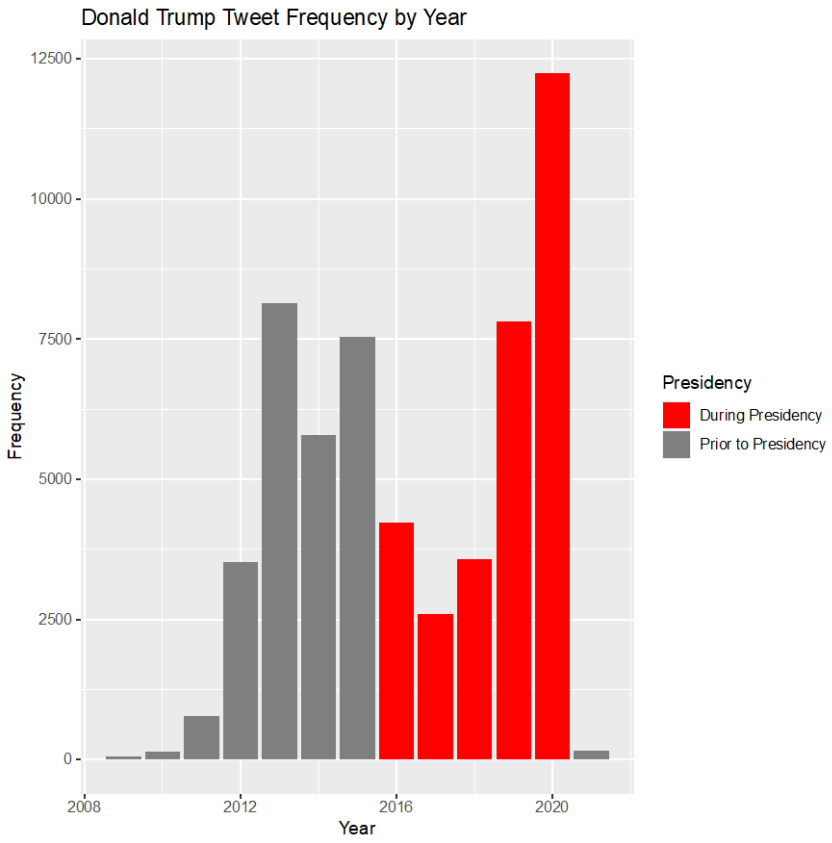


Figure 5. Donald Trump Tweets

With the goal of identifying Donald Trump's "Asian tweets" the first step was to identify which words Donald Trump was using to identify Asians (aside from "Asian"). Furthermore, I needed to see if there was a significant number of tweets about Asians to make the analysis viable.

By parsing and cleaning all of Donald Trump's tweets, I was able to tokenize the individual words in his tweets and plot them on the frequency chart shown in figure 6. Figure 6 shows that "china" is the 19th most-used term in his tweets. Using "china" as a keyword to identify tweets about Asians would provide a significant subset of tweets for the analysis. There are also tweets in which Donald Trump refers to Asians that contain keywords such as "asia" and "wuhan." The different variations of these words were also included as keywords.

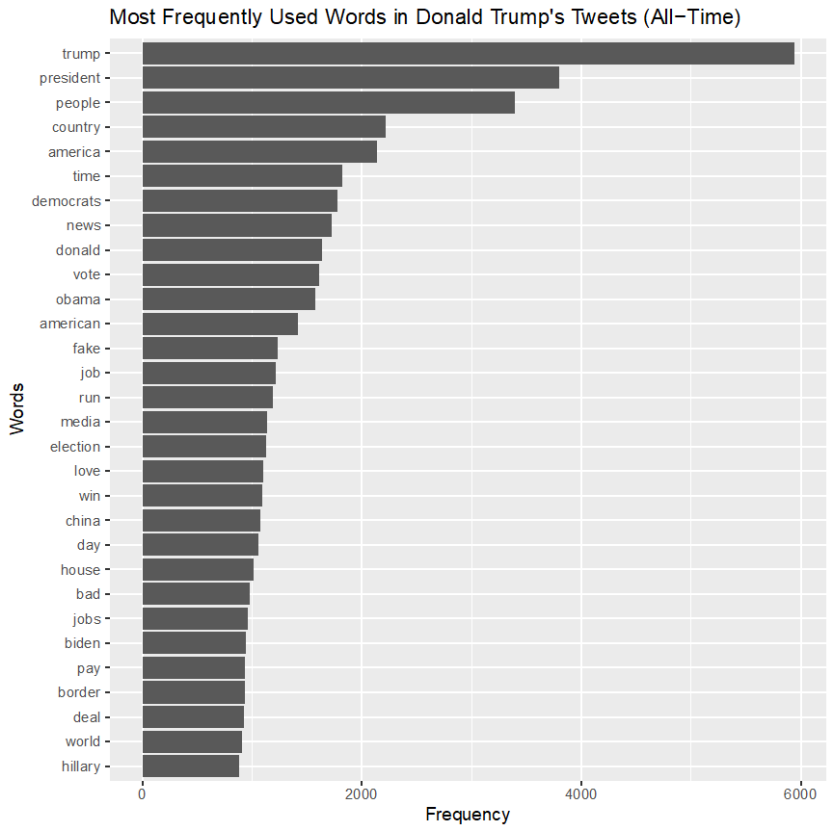
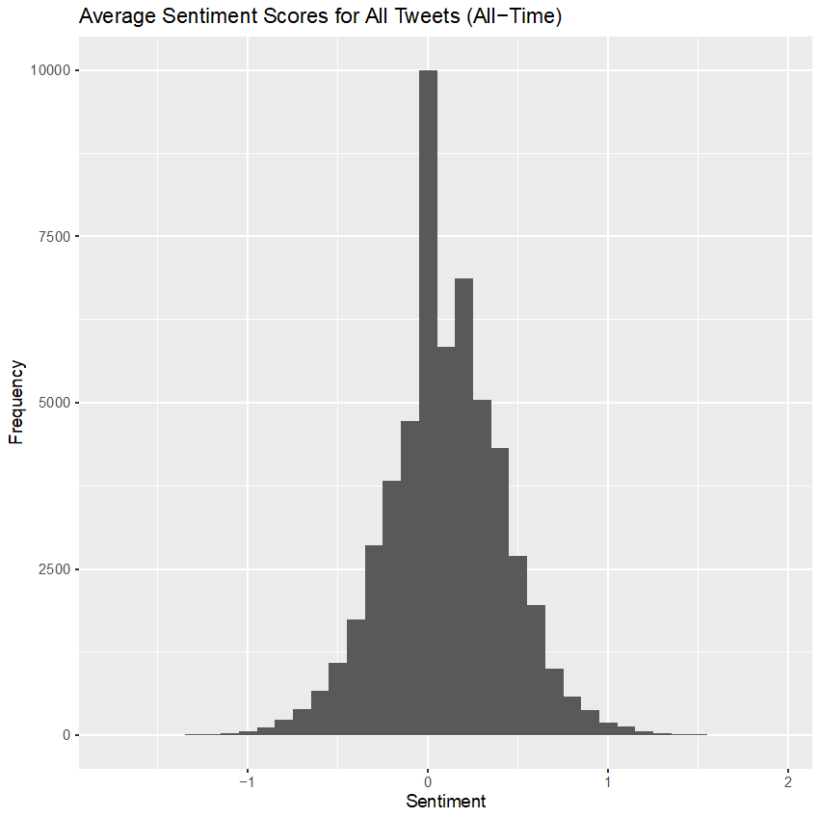


Figure 6. Frequently Used Words in Donald Trump's Tweets (All-Time)

The sentiment of each of Donald Trump's tweets were analyzed given a sentiment score. The sentiment was analyzed using the *AFINN* lexicon which assigns a score between -5 and 5 to each tweet. Negative scores indicate negative sentiment while positive scores indicate positive sentiment. The magnitude of the score indicates the degree by which a tweet's sentiment is negative or positive.

Figure 7 shows the distribution of the average sentiment scores for each of Donald Trump's tweets. We can see that a lot of his tweets are neutral (*sentiment = 0*). However, we can get a closer look at the distribution through the descriptive statistics of the graph.

From the descriptive statistics, we can see that the sentiments of his tweets are slightly positive judging from the central tendency. Furthermore, it seems like there is a large range of sentiment from his most positive tweets scoring 1.94 and his most negative tweets scoring -1.7.




```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
-1.69768 -0.08944 0.08030 0.09666 0.30000 1.93750
```

Figure 7. Distribution of Average Sentiment Scores for All of Donald Trump's Tweets

Figure 8 shows the distribution of the average sentiment scores for each of Donald Trump's tweets that refers to Asians. It is similar to figure 8 with regards to the pattern of distribution. It seems like Trump's tweets about Asians are largely neutral as well.

However, the descriptive statistics suggest that the central tendency is slightly lower than the distribution of all Trump's tweets; however, still slightly positive.

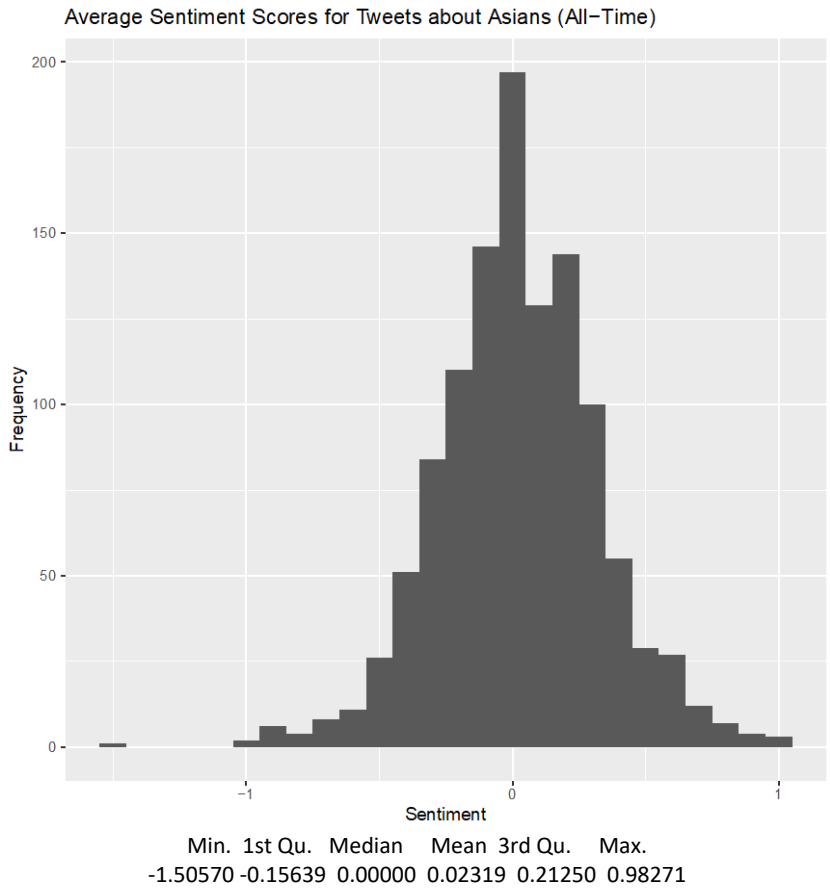


Figure 8. Distribution of Average Sentiment Scores for Donald Trump's Tweets about Asians (All-Time)

Figure 9 shows a comparison of the central tendencies (the mean) of the average sentiment scores of Donald Trump's tweets about Asians to those not about Asians through the difference of means test. The right side of the table shows the mean of the average sentiments of tweets not about Asians and the mean of average sentiments of tweets about Asians, respectively.

The comparison of means is statistically significant ($p < 0.05$). This means that we can reject the null hypothesis which states that the means are the same. Through this, we can conclude that Donald Trump's tweets that refer to Asians have a lower sentiment than tweets that do not refer to Asians.

Test statistic	df	P value	Alternative hypothesis	mean of x	mean of y
8.375	1215	1.511e-16 * * *	two.sided	0.09824	0.02319

Figure 9. t-test Between Donald Trump's Tweets About Asians and Not About Asians

Figure 10 tests sentiments as a predictor for the frequency of anti-Asian hate crimes. The test shows that Donald Trump's tweet sentiment is a statistically significant predictor with regards to the frequency of anti-Asian hate crimes. We can expect the frequency of anti-Asian hate crimes to increase by 0.02 as the sentiment increases by a factor of 1.

	nooffsetdf$frequency		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	0.26	0.25 – 0.27	**<0.001**
nooffsetdf$sentiment	0.02	0.00 – 0.04	**0.044**
nooffsetdf$asianTRUE	-0.03	-0.07 – 0.01	0.188
Observations	32786		
R^2 / R^2 adjusted	0.000 / 0.000		

Figure 10. OLS Model with No Leading Days

In reality, it is hard to believe that the effects of Donald Trump's negative sentiment tweets about Asians will move its way into the physical world that same day. Figure 11 runs the test for up to four days leading a tweet from Donald Trump.

Figure 11 shows that the only other statistically significant variable is in predicting the frequency of anti-Asian hate crimes is *asian* four days leading the initial tweet. We can expect that while holding the sentiment constant, that tweets about Asians will decrease the frequency of anti-Asian hate crimes.

	nooffsetdf$frequency			oneoffsetdf$frequency1			twooffsetdf$frequency2			threeoffsetdf$frequency3			fouroffsetdf$frequency4		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	0.26	0.25 – 0.27	**<0.001**	0.25	0.24 – 0.25	**<0.001**	0.25	0.25 – 0.26	**<0.001**	0.25	0.24 – 0.26	**<0.001**	0.26	0.25 – 0.27	**<0.001**
nooffsetdf$sentiment	0.02	0.00 – 0.04	**0.044**												
nooffsetdf$asianTRUE	-0.03	-0.07 – 0.01	0.188												
oneoffsetdf$sentiment				0.01	-0.00 – 0.03	0.137									
oneoffsetdf$asianTRUE				0.02	-0.02 – 0.06	0.239									
twooffsetdf$sentiment							-0.01	-0.03 – 0.01	0.400						
twooffsetdf$asianTRUE							-0.00	-0.04 – 0.04	0.851						
threeoffsetdf$sentiment										-0.01	-0.03 – 0.01	0.384			
threeoffsetdf$asianTRUE										-0.02	-0.06 – 0.02	0.452			
fouroffsetdf$sentiment													0.02	-0.00 – 0.03	0.102
fouroffsetdf$asianTRUE													-0.05	-0.09 – -0.01	**0.019**
Observations	32786			32786			32786			32786			32786		
R^2 / R^2 adjusted	0.000 / 0.000			0.000 / 0.000			0.000 / -0.000			0.000 / -0.000			0.000 / 0.000		

Summary and Discussion

Although not conclusive, my analysis suggests that my theory can be rejected. There exist significant predictors for the frequency of anti-Asian hate crimes; however, it is not the relationship that my initial theory had stated. If my analysis were consistent with my theory, an increase in sentiment would indicate a decrease in the frequency of anti-Asian hate crimes. However, my analysis does suggest that Donald Trump tweets about Asians are largely negative in sentiment than those that are not.

The preliminary analysis still suggests that Trump's negative tweets about Asians may be a causal mechanism for the rise in anti-Asian hate crimes. The limitations of the UCR program's hate crime database may be part of the inconsistency between my theory and analysis. This database relies solely on those crimes that are reported to the FBI. The studies that I reviewed in

the overview use the experimental method of surveying which may indicate that a large portion of hate crimes do not get reported to the FBI.