Trump Effect: Hate Speech on Twitter and Hate Crime

Anina Sophie Jauris, Vanisa Vongphanakhone, Dung Duc Huynh

6/13/2020

```r
hate_crime <- read_csv("../raw-data/hate_crime.csv")
```

```
## Parsed with column specification:
## cols(
##   .default = col_character(),
##   INCIDENT_ID = col_double(),
##   DATA_YEAR = col_double(),
##   ADULT_VICTIM_COUNT = col_logical(),
##   JUVENILE_VICTIM_COUNT = col_logical(),
##   TOTAL_OFFENDER_COUNT = col_double(),
##   ADULT_OFFENDER_COUNT = col_logical(),
##   JUVENILE_OFFENDER_COUNT = col_logical(),
##   OFFENDER_ETHNICITY = col_logical(),
##   VICTIM_COUNT = col_double(),
##   TOTAL_INDIVIDUAL_VICTIMS = col_double()
## )

## See spec(...) for full column specifications.

## Warning: 22965 parsing failures.
##    row                     col           expected                 actual                         fil
## 100969 JUVENILE_OFFENDER_COUNT 1/0/T/F/TRUE/FALSE 2                      '../raw-data/hate_crime.csv
## 100969 OFFENDER_ETHNICITY      1/0/T/F/TRUE/FALSE Not Hispanic or Latino '../raw-data/hate_crime.csv
## 123891 OFFENDER_ETHNICITY      1/0/T/F/TRUE/FALSE Unknown                '../raw-data/hate_crime.csv
## 140736 JUVENILE_OFFENDER_COUNT 1/0/T/F/TRUE/FALSE 2                      '../raw-data/hate_crime.csv
## 140736 OFFENDER_ETHNICITY      1/0/T/F/TRUE/FALSE Unknown                '../raw-data/hate_crime.csv
## ...... ....................... .................. ...................... ...........................
## See problems(...) for more details.
```

```r
data_HateCrimeTrump <- read_csv("../processed-data/data_HateCrimeTrump.csv")
```

```
## Parsed with column specification:
## cols(
##   .default = col_double(),
##   state = col_character(),
##   state_name = col_character()
## )
## See spec(...) for full column specifications.
```

```r
twitter_usage <- read_csv("../raw-data/twitter_usage_yearly.csv")
```

```
## Parsed with column specification:
## cols(
##   year = col_double(),
##   state = col_character(),
```

```
##   total = col_double()
## )
```

```r
freq_neg_trump_hispanic <- read_csv("../processed-data/freq_neg_trump_hispanic.csv")
```

```
## Warning: Missing column names filled in: 'X1' [1]

## Parsed with column specification:
## cols(
##   X1 = col_double(),
##   week = col_double(),
##   freq = col_double(),
##   total_negative = col_double(),
##   total_tweet = col_double()
## )
```

```r
freq_neg_trump_chinese <- read_csv("../processed-data/freq_neg_trump_chinese.csv")
```

```
## Warning: Missing column names filled in: 'X1' [1]

## Parsed with column specification:
## cols(
##   X1 = col_double(),
##   week = col_double(),
##   freq = col_double(),
##   total_negative = col_double(),
##   total_tweet = col_double()
## )
```

```r
usa_tweets_hispanic <- read_csv("../processed-data/usa_tweets_hispanic.csv")
```

```
## Parsed with column specification:
## cols(
##   date = col_date(format = ""),
##   text_clean = col_character(),
##   WordCount = col_double(),
##   SentimentGI = col_double(),
##   state = col_character()
## )
```

```r
usa_tweets_coronavirus <- read_csv("../processed-data/usa_tweets_coronavirus.csv")
```

```
## Parsed with column specification:
## cols(
##   date = col_date(format = ""),
##   text_clean = col_character(),
##   WordCount = col_double(),
##   SentimentGI = col_double(),
##   state = col_character()
## )
```

Introduction

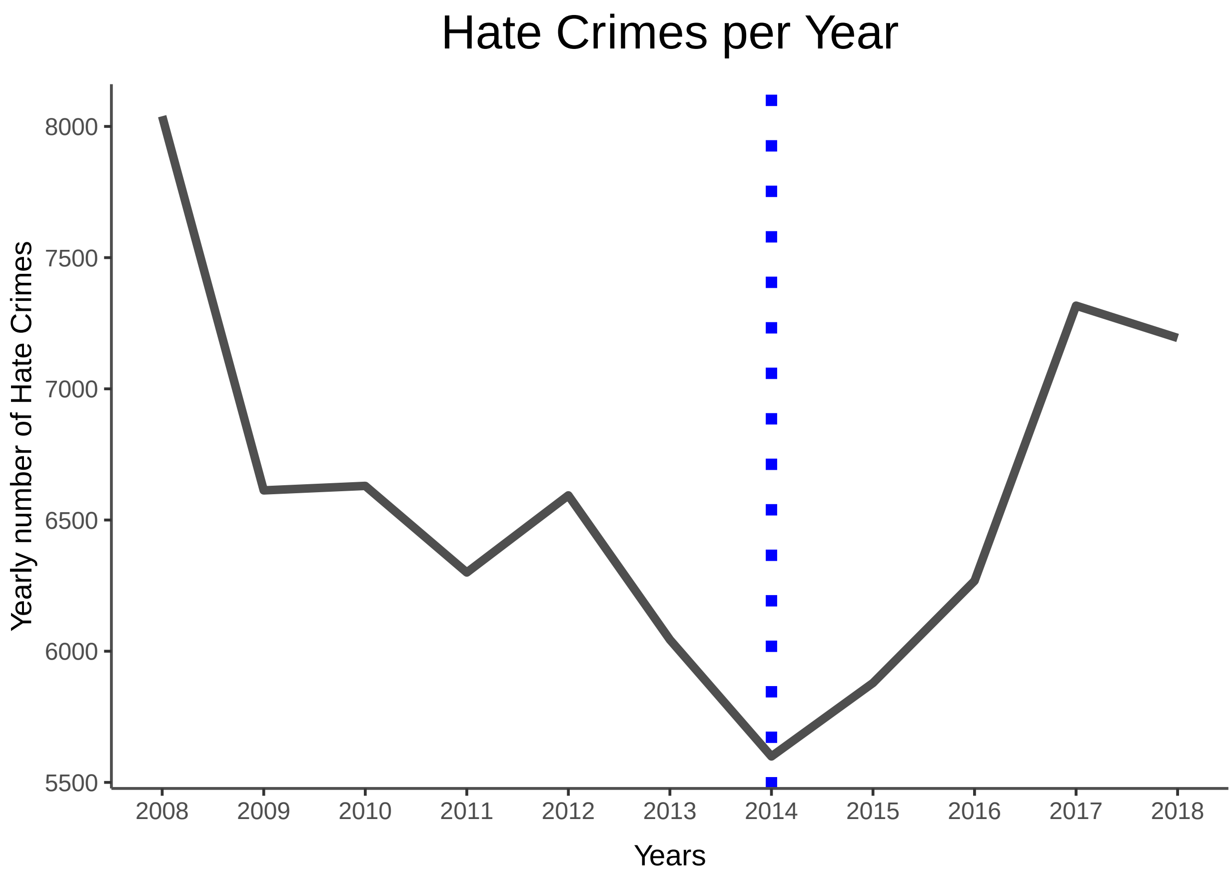
Over the last two decades the use of social media platforms like Facebook and Twitter became an important part of people's lives. Those platforms are not only used for social interaction, they also turned into a major source of information. Compared to other social media platforms Twitter stands out in particular, as 71% of all U.S. users get news on the site (2019a (n.d.)). A large share of news content is directly provided by political figures, with Donald Trump probably being the most active political Twitter user. This direct way of spreading information and opinions combines several advantages. For one thing, the reach and speed of news on Twitter is high, making it possible to instantly react to recent events. Secondly, social platforms are allowing biased statements, enabling politicians to promote their own agenda.

With the beginning of his presidential campaign in 2015, Trump started to regularly denounce several minority groups in order to support his arguments. Some of his tweets were highly prejudiced and often controversially discussed in the context of acceptable speech. Twitter's chief executive, Jack Dorsey, stated in an interview in 2019 that even though some statements of political figures like Trump are in violation of the terms and conditions of Twitter, they are of public interest and therefore remain unaffected by regulations even if they are promoting hate speech (Rogan (2019)). Hate speech broadly refers to every expression of hatred towards a stigmatized group, directed on specific characteristics like race, religion, ethnicity, etc (Álvarez-Benjumea and Winter (2018); Titley, Keen, and Földi (2014)). This immunity remained almost unchanged over the course of Trump's presidency, with a few recent exceptions in the end of May[1]. This shows that Trump had a privileged position in online environments without having to face any consequences for his prejudiced tweets.

This lack of regulation and Trump's elite status raised the question of a potential link between Trumps Twitter use and hate speech and hate crime, respectively. Yearly hate crime statistics provided by the FBI suggest that overall hate crimes in the U.S. have increased since the year **of 2014 (We need to make a table). This rising trend was also seen for hate crimes with a Mexican/ Hispanic bias, starting from 2015** . Increasing hate crimes therefore coincide with president Trumps progressive use of Twitter for political communication. Müller and Schwarz (2019) made an effort to identify a causal link between Trump's prejudiced tweets towards Muslims and anti-Islamic motivated crime and found a significant effect.

```r
hate_crime %>%
  group_by(DATA_YEAR) %>%
  filter_at(vars(starts_with("DATA_YE")), all_vars(. > 2007)) %>%
  summarise(incident_num = n()) %>%
  ggplot(aes(x = DATA_YEAR, y = incident_num))+
  geom_line(size=1.5, col="grey31") +
  ggtitle("Hate Crimes per Year") +
  scale_x_continuous("Years", breaks = seq(2007, 2018, 1))+
  scale_y_continuous("Yearly number of Hate Crimes") +
  theme( plot.title = element_text(size= 18, vjust=2, hjust=0.5),
         axis.title.x = element_text(vjust=-1),
         axis.line = element_line(color = "grey31"),
         panel.background = element_blank()) +
  geom_vline(xintercept  = 2014, linetype="dotted",
             color = "blue", size=2)
```

[1]On May 26th Twitter appended fact-checking labels to misleading tweets about mail-in voting. On May 29th Twitter again took action and labelled one of Trump's tweets as "glorifying violence" in which he threatened to shoot people if looting happens. The most recent intervention was the removal of a campaign video because of copyright complaints on June 3rd. In response, Trump issued an executive order to strip Twitter's legal protection to make them liable for the content on their platform (Conger and 2020 (n.d.))

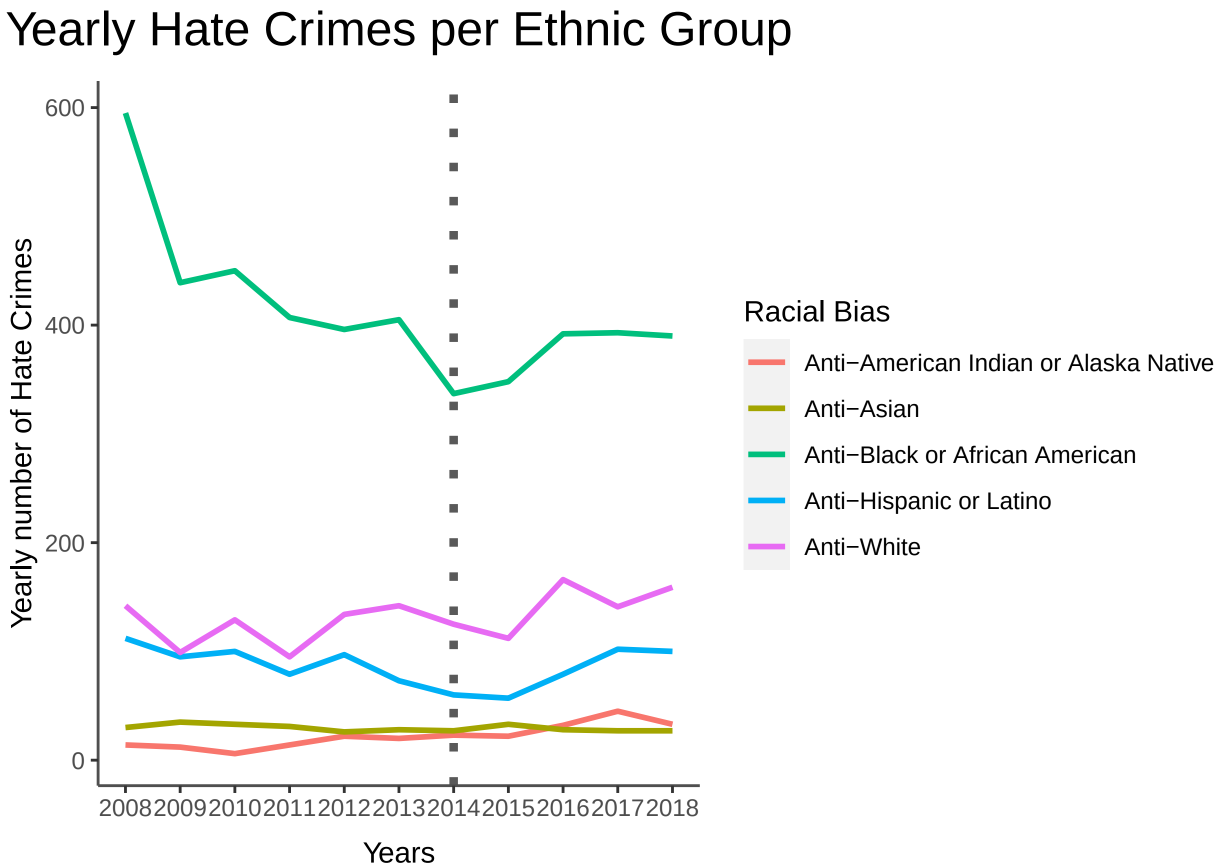


In this paper we are also interested in uncovering a causal link between Trump's Tweets and hate speech on Twitter on the one hand and Trump's Tweets and hate crime on the other **(maybe make a graph to visualize but not necessary .. only when we have time :P)**. We will focus on two different minority groups that both used to be Trump's main target group at one point in time. The first minority group is the population of Mexicans and Hispanics, who he addressed especially in the beginning of his presidential campaign. He often referred to them as illegal aliens, associating them with crime, violence and drug trafficking. For this group we have data on Trump tweets, tweets using the same keywords as Trump and data from hate crime statistics which enables us to draw a connection between both suggested links respectively.

```
hate_crime %>%
  filter_at(vars(starts_with("DATA_YE")), all_vars(. > 2007)) %>%
  mutate(BIAS_DESC = strsplit(as.character(BIAS_DESC), ";")) %>%
  unnest(BIAS_DESC) %>%
  filter(BIAS_DESC==c("Anti-White", "Anti-Black or African American",
                      "Anti-Asian", "Anti-American Indian or Alaska Native",
                      "Anti-Hispanic or Latino")) %>%
  group_by(DATA_YEAR, BIAS_DESC) %>%
  summarise(incident_num = n()) %>%
  ggplot(aes(x = DATA_YEAR, y = incident_num)) +
  geom_line(aes(col = BIAS_DESC), size=1) +
  ggtitle("Yearly Hate Crimes per Ethnic Group") +
  labs(col = "Racial Bias") +
  scale_x_continuous("Years", breaks = seq(2007, 2018, 1))+
  scale_y_continuous("Yearly number of Hate Crimes") +
  theme( plot.title = element_text(size= 18, vjust=2, hjust=0.5),
         axis.title.x = element_text(vjust=-1),
```

```
        axis.line = element_line(color = "grey31"),
        panel.background = element_blank()) +
  geom_vline(xintercept  = 2014, linetype="dotted",
            color = "grey35", size=1.5)
```

```
## Warning in BIAS_DESC == c("Anti-White", "Anti-Black or African American", :
## longer object length is not a multiple of shorter object length
```



The second minority group is the group of Chinese people. Starting with the Corona Virus pandemic Trump made Chinese people responsible for the health issues in the U.S.. Since this are rather new developments, we only have Trump's tweets and tweets using the same keywords as Trump. Data on hate crimes towards Chinese people is not available yet, therefore we can only look at the first suggested link. To see if the Chinese minority group in the U.S. is about to face the same increase in hate crime incidences as the Hispanic minority group, we will compare the hate speech outcomes of the two groups. By finding similar or dissimilar results we might be able to make a conclusion for future trends of hate crimes towards Chinese.

We therefore try to answer three different questions. If Trump's online rhetoric is influencing other people's online expression, if Trump's tweets have an effect on hate crime and if Trump's tweets about Chinese might result in more future hate crime towards Chinese.

Theory & Hypotheses

People care to a large extent about what other people think of them. To navigate what behavior is socially accepted, people usually adjust their actions to social norms (Asch (1975); Cialdini and Goldstein (2004)). This norm conforming behavior prevents people from social exclusion and from being sanctioned. Over the last decades there has been a clear social norm against public expression of hate (Schaffner (2018)). Social norms therefore were a main motivation to hide one's prejudice and to not engage in hate speech or hate crime (Crandall, Eshleman, and O'brien (2002)). However, recent trends imply that there has been a disruption of these norms, leading to a greater acceptability of explicitly hateful and degrading behavior.

To learn about social norms people rely on cues from others to infer "normal behavior" (Zitek and Hebl (2007); Schaffner (2018)). These cues are often taken from observations of how other people act (Bicchieri and Xiao (2009); Krupka and Weber (2008)). Especially influential in these processes are people of public interest. Due to his social status as president of the United States and his large news coverage in online and offline media, Trump is likely to be one of the key figures shaping social norms. He might be seen as reference point to whom people adjust their behavior. When Trump uses hateful and clearly racist rhetoric the outer bounds of acceptable speech will expand. People with prejudiced thoughts who formerly suppressed hateful actions might now adjust their behavior to the new limits, resulting in more expression of prejudice. This could be either in the form of online commenting or in the form of physically or verbally assaulting people in real life.

This tendency was also found in prior experimental research. Gervais (2014) was able to show that, republicans are especially prone to negative commenting. When being exposed to like-minded uncivil political rhetoric, republicans were more likely to copy uncivil language in their own messages compared to democrats. Another experiment by Schaffner (2016) was focusing directly on the effect of tweets made by Trump. Individuals in the treatment group, who saw his negative tweets about Muslims, were significantly more likely to also engage in hate speech towards Muslims. Trump's remarks therefore served as a cue, constituting the bounds of acceptable speech and legitimizing uncivil talk. We argue that by creating an environment of higher tolerance for publicly displaying hate, this will not only show in online contexts but also transfer to real life behavior. This might result in open discrimination or even hate crime.

One of the reasons for Trump's considerable influence might be the existence of a general norm of authority compliance, even in the absence of strong incentives (Karakostas and Zizzo (2016)). The status as president might also make him thought of as more prestigious, more knowledgeable and more powerful. People might anticipate Trump to have greater access to crucial political information, unavailable to regular civilians, that informs him in making political statements. This expected information asymmetry might lead some people to think of his statements as more legitimate and trustworthy, regardless of whether this is true or not. Trump might also be seen as "voice of society". Since he won the 2016 elections, he resembles an aggregate of individuals opinion (Bursztyn, Egorov, and Fiorin (2017)). His electoral win thereby informed people about the opinions of the people surrounding them, suggesting that a large share of society actually approves of Trump's negative attitudes towards certain minority groups. This makes sanctioning or social exclusion for publicly expressing prejudice less likely and could therefore cause an increase in hate speech and hate crime.

Literature on framing and priming suggests, that politicians often use framing strategies to systematically manipulate public opinion (Schaffner and Sellers (2009); Chong and Druckman (2007)). By emphasizing certain aspects and leaving out others, opinions on controversial topics, can be modified to fall in line with the opinion held by the political actor. Likewise, Trump seems to use Twitter as political instrument, intentionally placing negative statements to strengthen his own agenda.

One of many examples is president Trump directly relating Mexican immigrants to higher crime rates, drugs, human trafficking and other bad behaviors, to strengthen his anti-immigration politics[2]. He often instrumentalized single incidents and portrayed them as being representative for the minority group as a whole. By ascribing bad characteristics and highlighting the distinctiveness between ingroup and outgroup, he publicly defined group positions claiming that one group is superior to the other. This insinuated superiority and distinctiveness might elicit feelings of aversion or legitimize already existing aversion (Blumer (1958)). It might also unconsciously strengthen peoples' ingroup identification and trigger a feeling of threat and anger towards the outgroup members (Blumer (1958); Gervais (2013)). This threat is likely to be provoked by the supposed "bad" behavior or "bad" consequences that come with the outgroup members. The minority group of Mexicans might e.g. be perceived as an economic threat or a threat to inner security, since Trump often related

[2]On Oct 19, 2016 08:22:19 pm he posted for example „Druggies, drug dealers, rapists and killers are coming across the southern border. When will the U.S. get smart and stop this travesty?". On Apr 25, 2017 07:36:28 am he tweeted "Don't let the fake media tell you that I have changed my position on the WALL. It will get built and help stop drugs, human trafficking etc.". These are just some examples of his frequently tweeted aversion towards Mexican immigrants throughout his presidential campaign and presidency.

them to crime and taking over jobs. When Trump used Terms like "invasion of illegal immigrants", these threats might feel even more severe. Given the current circumstances, with the Covid-19-pandemic, Chinese people on the other hand might have be perceived as a health threat. When the first consequences started to show in the U.S., Trump held China responsible, accusing them of purposely withholding information or even purposely spreading the virus. He regularly called the Covid-19-virus "Chinese Virus" or "China Virus". But even if Chinese people were not primarily seen as health threat themselves, people might have adopted Trump's argumentation, making Chinese people responsible for the critical economic and health situation in the U.S.

What might have intensified these mechanisms even further is the "sorting" mechanism in online environments. People tend to self-select into online platforms or groups used by like-minded people which leads to the formation of so-called echo chambers (Wahlström and Törnberg (2019); Gervais (2013)). By hearing or reading like-minded statements, peoples' beliefs are consolidated which might further increase polarization of opinions. As an analysis of the Pew Research Center (2019b (n.d.)) has shown, Trump's Twitter follower are more likely to be republican or republican-leaning. Thus, one can suggest that his followers already hold more conservative feelings and are more likely to be affected by Trumps tweets. However, tweets sent by the president probably reach a much larger share of people through either retweets, quoting tweets or media coverage. Therefore, a lot more Americans are exposed to Trump's tweets even if they don't follow him. That is why we are interested in the effect of Trump on twitter users in general and not on his followers in particular.

The above argumentation leads us to the following hypothesis.

- H1: If the weekly number of negative Trump tweets increases, the overall number of tweets concerning the same topics will increase.
- H2: If the weekly number of negative Trump tweets increases, the attitudes of tweets concerning the same topics will get more negative.
- H3: If the weekly number of negative Trump tweets increases, the weekly number of hate crime incidents will increase as well.

The first hypothesis looks at how prominent Trump tweets are in guiding public discussion and setting topics of interest, whereas the second hypothesis considers the negativity of the responding tweets. These hypotheses will be examined for the group of Mexicans and Chinese separately. The third hypothesis is focusing on hate crime. Since we only have data for hate crime towards Mexicans our analysis will restricted to this group only.

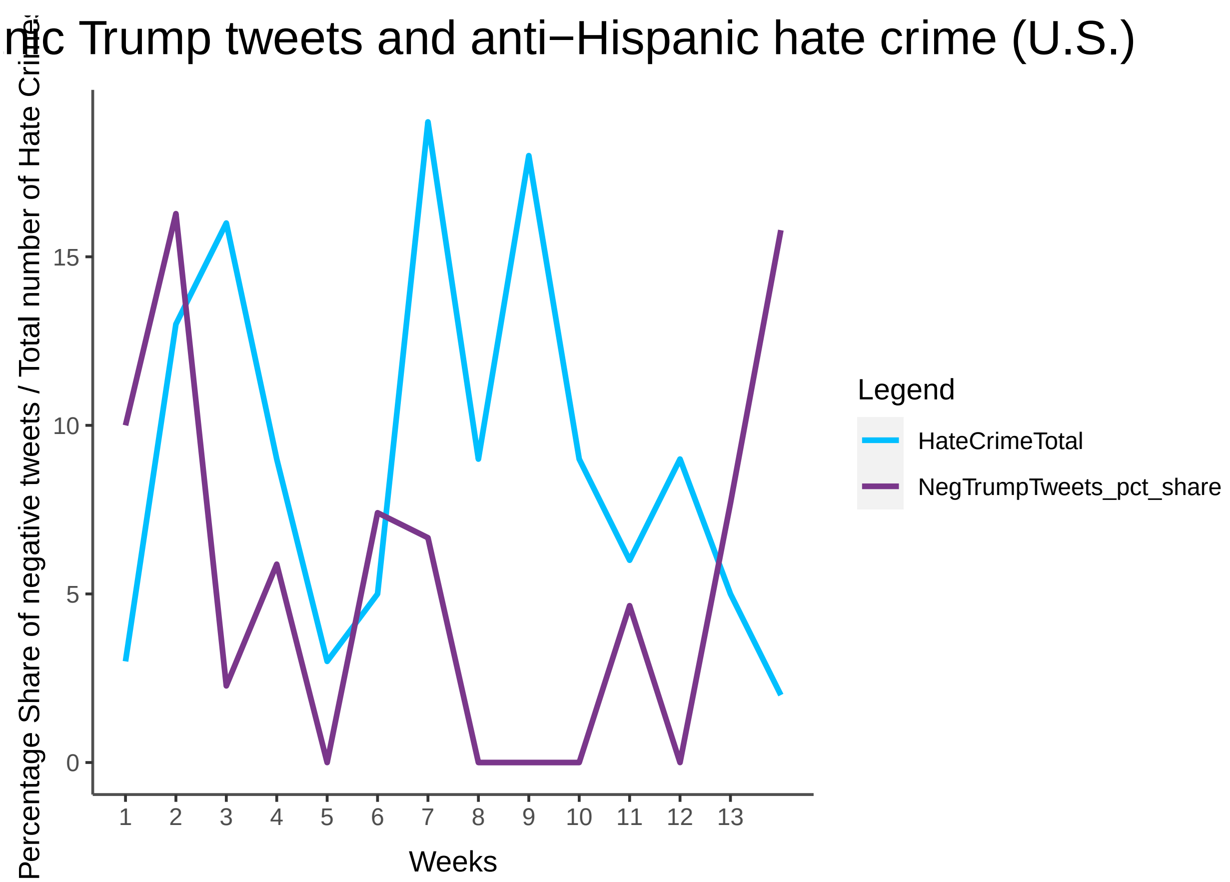
Data

We restricted the data collection for the dependent and explanatory variables on two particular timeframes. For Trump tweets, hate speech and hate crime towards Mexicans we set the timeframe to the 20th of January 2017 till the 20th of April 2017. For Chinese our timeframe was the 16th of March 2020 till the 10th of June 2020. The decision for those timeframes will be discussed below. To achieve a higher comparability between the trends towards Mexicans and Chinese we only looked at data since Trump officially became president. His social influence before and after his electoral victory might have been different. When being in office Trump was not only a very well-known public figure, he was also holding judicial power which gave him additional legal legitimacy. We therefore set the starting date of our data collection for Mexican tweets and hate crime to the day of his inauguration, the 20th of January 2017, even though he started commenting negatively a couple of months before. Roughly around the start of his presidential campaign. For the tweets against Chinese we set the starting date of data collection to the first day Trump mentioned the term "Chinese Virus" on Twitter. Since the Covid-19 pandemic is a relatively new topic we were only able to collect data until the 10th of June 2020. To have a similar timespan for both minority groups we also restricted the data collection for Mexicans to a time span of three months. Because the timespan is relatively short, we focus on weekly changes in our analysis.

For an overview over all collected variables and sources please have a look **at Table XY.**

Explanatory variable – Frequency of negative Trump tweets

```
## Anti-Hispanic hate crime (whole US)
data_HateCrimeTrump %>%
  group_by(week) %>%
  summarise(NegTrumpTweets_pct_share = sum(freq)*100/n(),
            HateCrimeTotal = sum(incident_number)) %>%
  gather(key = "variable", value = "value", -week) %>%
  ggplot(aes(x = week, y = value)) +
  geom_line(aes(color = variable), size=1) +
  scale_color_manual(values = c("deep sky blue", "MediumOrchid4")) +
  labs(title="Anti-Hispanic Trump tweets and anti-Hispanic hate crime (U.S.)") +
  labs(col = "Legend") +
  scale_x_continuous("Weeks", breaks = seq(1, 13, 1))+
  scale_y_continuous("Percentage Share of negative tweets / Total number of Hate Crimes") +
  theme( plot.title = element_text(size= 18, vjust=2, hjust=0.5),
         axis.title.x = element_text(vjust=-1),
         axis.line = element_line(color = "grey31"),
         panel.background = element_blank())
```

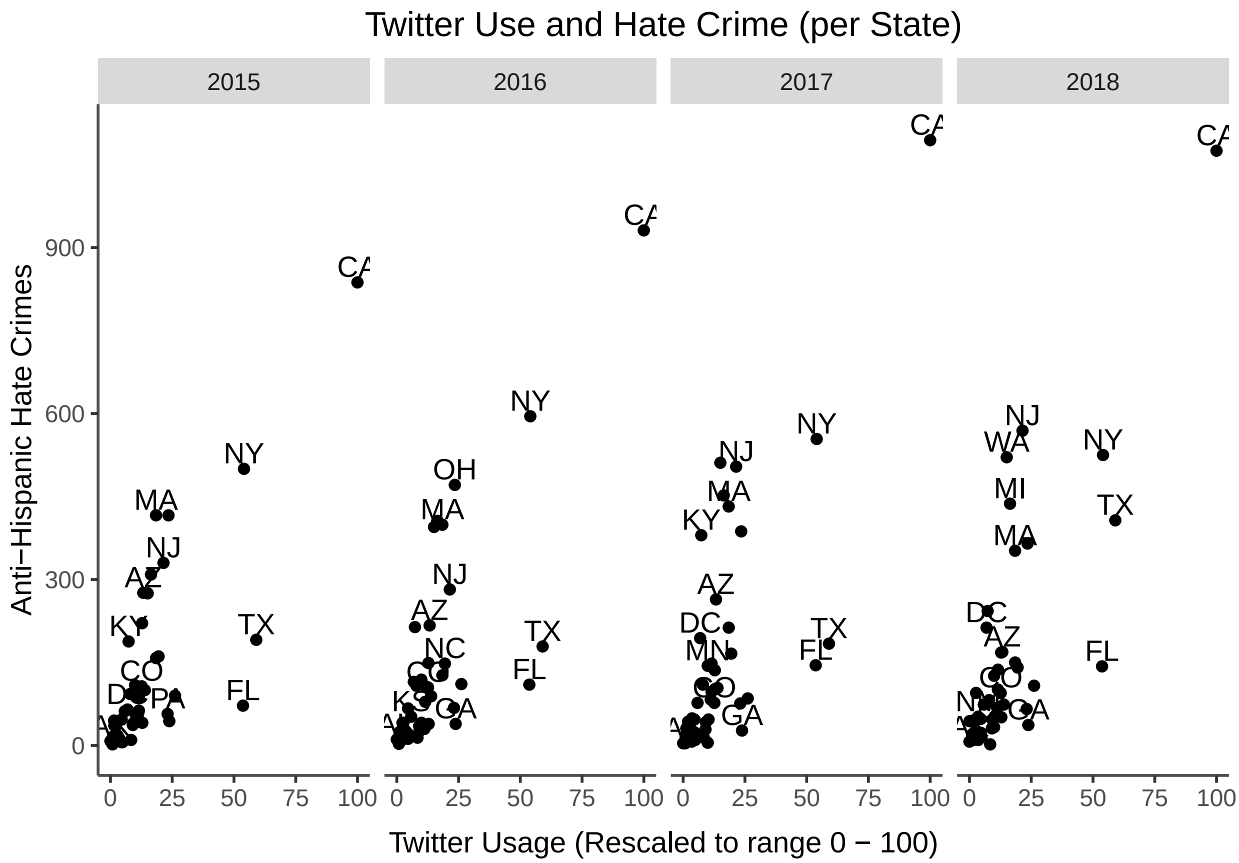



```
twitter_usage_state <- twitter_usage %>%
  group_by(state) %>%
  summarise(mean_usage14_15 = sum(total)/n()) # take the mean hashtag counts for years 2014 and 2015

twitter_usage_state$mean_usage14_15 <-
  scales::rescale(twitter_usage_state$mean_usage14_15, to=c(0,100)) # rescaling
```

```r
hate_crime %>%
  group_by(DATA_YEAR, STATE_ABBR) %>%
  summarise(incident_num = n()) %>%
  left_join(twitter_usage_state, by=c('STATE_ABBR' = 'state')) %>%
  filter(2015 <= DATA_YEAR & DATA_YEAR <= 2018) %>%
  ggplot(aes(x=mean_usage14_15, y=incident_num)) +
  geom_point() +
  geom_text(aes(label= STATE_ABBR), vjust = 0, nudge_y = 10, check_overlap = T) +
  labs(title="Twitter Use and Hate Crime (per State)", x="Twitter Usage (Rescaled to range 0 - 100)", y=
  labs(col = "State")+
  theme( plot.title = element_text( hjust=0.5),
         axis.title.x = element_text(vjust=-1),
         axis.line = element_line(color = "grey31"),
         panel.background = element_blank()) +
  facet_grid(. ~ DATA_YEAR)
```

```
## Warning: Removed 6 rows containing missing values (geom_point).
```

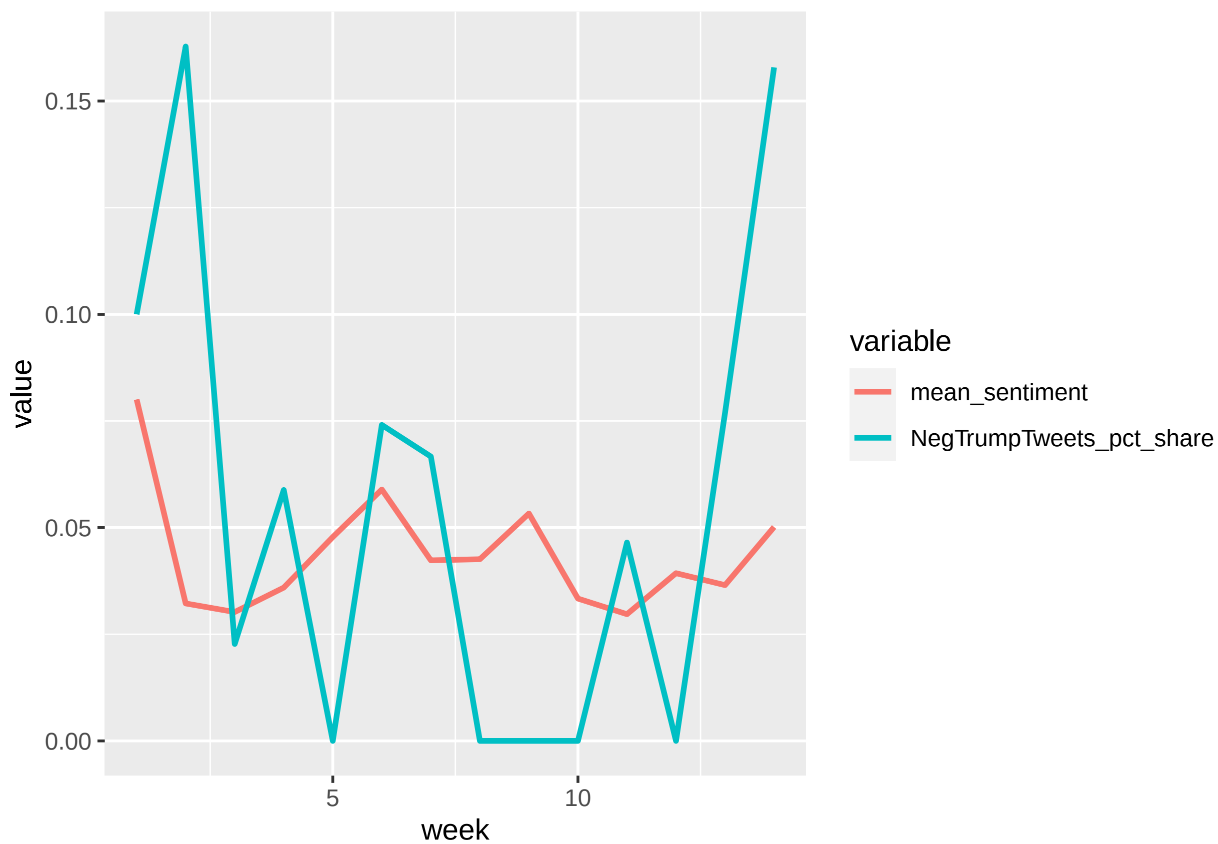
```
## Warning: Removed 6 rows containing missing values (geom_text).
```




```r
freq_trump_hispanic <- freq_neg_trump_hispanic
freq_trump_hispanic$state <- "CA"
for (state in unique(usa_tweets_hispanic$state)){
  if (state != "CA"){
    tmp <- freq_neg_trump_hispanic
```

```
    tmp$state <- state
    freq_trump_hispanic <- rbind(freq_trump_hispanic, tmp)
  }
}



freq_trump_hispanic %>%
  left_join(usa_tweets_hispanic %>%
              mutate(week = cut.Date(date, breaks = "1 week", labels = FALSE)) %>%
              arrange(date, week) %>%
              group_by(week, state) %>%
              summarise(sentiment = mean(SentimentGI))) %>%
  group_by(week) %>%
  summarise(NegTrumpTweets_pct_share = sum(freq)/n(),
            mean_sentiment = mean(sentiment, na.rm = TRUE)) %>%
  gather(key = "variable", value = "value", -week) %>%
  ggplot(aes(x = week, y = value)) +
  geom_line(aes(color = variable), size=1)
```

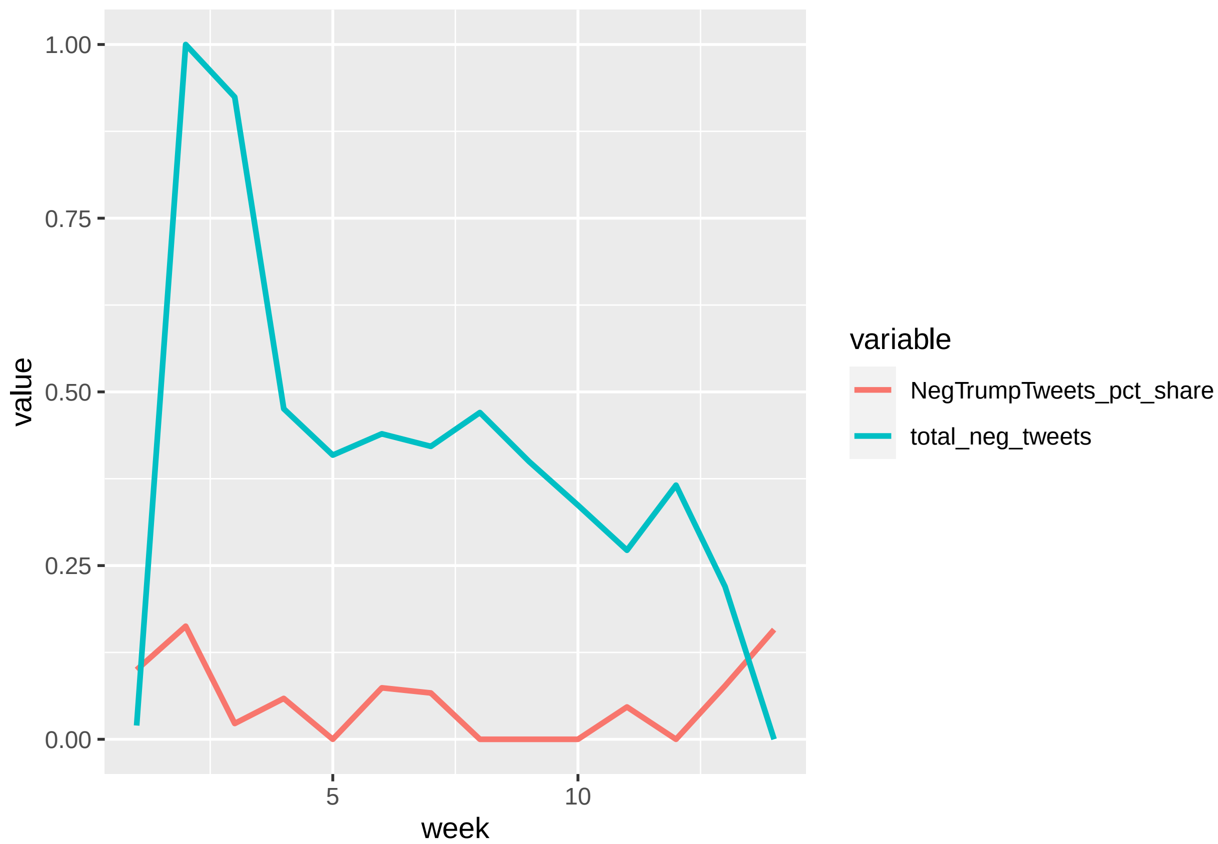
```
## Joining, by = c("week", "state")
```




```
freq_trump_hispanic %>%
  left_join(usa_tweets_hispanic %>%
              mutate(week = cut.Date(date, breaks = "1 week", labels = FALSE)) %>%
              arrange(date, week) %>%
```

```
              group_by(week, state) %>%
              summarise(negative_tweets = sum(SentimentGI < 0))) %>%
  group_by(week) %>%
  summarise(NegTrumpTweets_pct_share = sum(freq)/n(),
            total_neg_tweets = sum(negative_tweets, na.rm = TRUE)) %>%
  mutate(total_neg_tweets = (total_neg_tweets - min(total_neg_tweets))/ (max(total_neg_tweets) - min(to
  gather(key = "variable", value = "value", -week) %>%
  ggplot(aes(x = week, y = value)) +
  geom_line(aes(color = variable), size=1)
```

```
## Joining, by = c("week", "state")
```

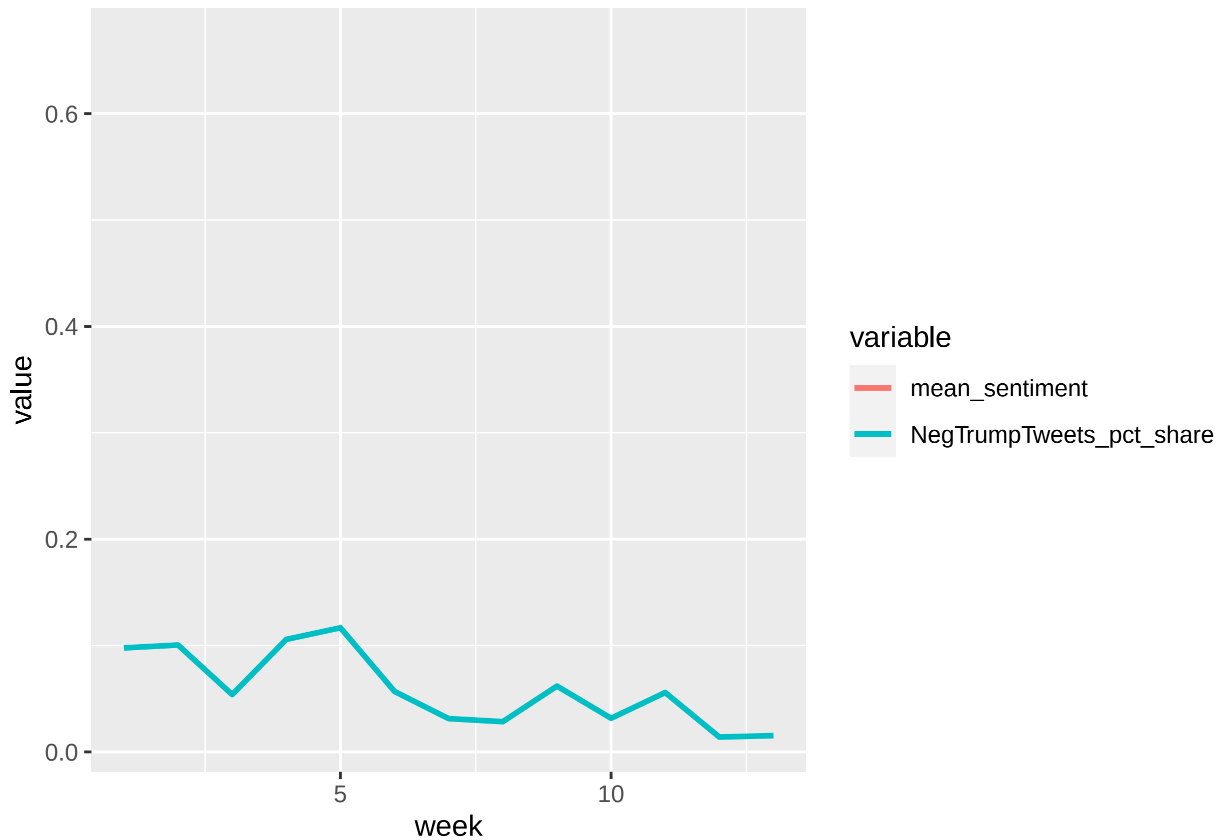



```
freq_trump_chinese <- freq_neg_trump_chinese
freq_trump_chinese$state <- "CA"
for (state in unique(usa_tweets_hispanic$state)){
  if (state != "CA"){
    tmp <- freq_neg_trump_chinese
    tmp$state <- state
    freq_trump_chinese <- rbind(freq_trump_chinese, tmp)
  }
}



freq_trump_chinese %>%
  left_join(usa_tweets_coronavirus %>%
```

```
              mutate(week = cut.Date(date, breaks = "1 week", labels = FALSE)) %>%
              arrange(date, week) %>%
              group_by(week, state) %>%
              summarise(sentiment = mean(SentimentGI))) %>%
  group_by(week) %>%
  summarise(NegTrumpTweets_pct_share = sum(freq)/n(),
            mean_sentiment = mean(sentiment, na.rm = TRUE)) %>%
  gather(key = "variable", value = "value", -week) %>%
  ggplot(aes(x = week, y = value)) +
  geom_line(aes(color = variable), size=1)
```

```
## Joining, by = c("week", "state")
```

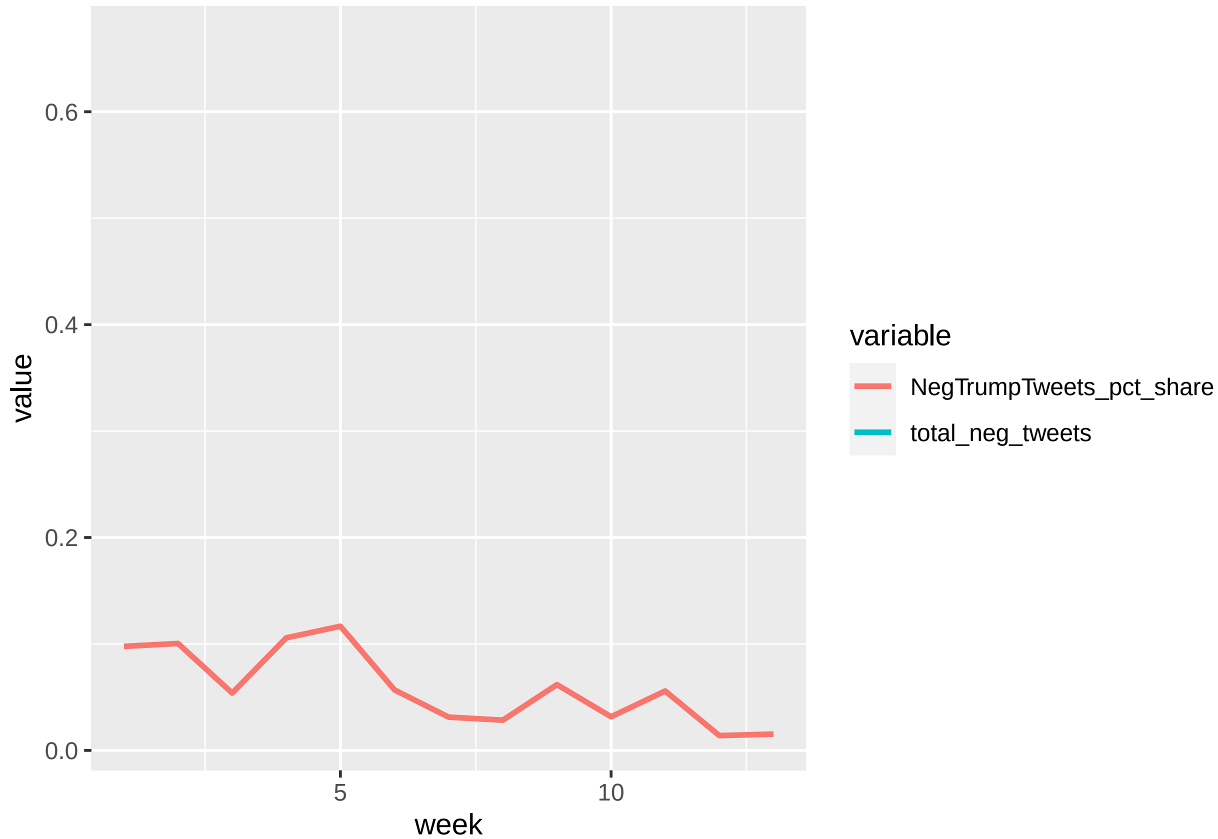
```
## Warning: Removed 15 row(s) containing missing values (geom_path).
```




```
freq_trump_chinese %>%
  left_join(usa_tweets_coronavirus %>%
              mutate(week = cut.Date(date, breaks = "1 week", labels = FALSE)) %>%
              arrange(date, week) %>%
              group_by(week, state) %>%
              summarise(negative_tweets = sum(SentimentGI < 0))) %>%
  group_by(week) %>%
  summarise(NegTrumpTweets_pct_share = sum(freq)/n(),
            total_neg_tweets = sum(negative_tweets, na.rm = TRUE)) %>%
  mutate(total_neg_tweets = (total_neg_tweets - min(total_neg_tweets))/ (max(total_neg_tweets) - min(to
  gather(key = "variable", value = "value", -week) %>%
  ggplot(aes(x = week, y = value)) +
```

```r
  geom_line(aes(color = variable), size=1)
```

```
## Joining, by = c("week", "state")
## Warning: Removed 15 row(s) containing missing values (geom_path).
```



For collecting Trump tweets we rely on the Trump Twitter Archive which was made available by Brendan Brown. The Archive covers almost every tweet ever made by Trump and is updated in real time. The website provides an advantage over extracting Trump's tweets directly from Twitter as it archives the tweets that Trump deletes on his Twitter account as well.

For our analysis we are interested in the frequency of negative tweets. To be able to calculate how frequently Trump was using hate speech towards a specific group we collected the number of negative Trump tweets and the overall number of Trump tweets per week and divided them.

E.g:

For the count of negative tweets we used a set of keywords which are likely to be related to the respective minority group. For details please see **Appendix Table: Keyword table**. We looked at every tweet containing one of these keywords separately and made sure that they were addressing Mexicans / Chinese. The categorization into positive and negative tweets was also done by hand. The overall weekly number of Trump tweets was taken from the Twitter archive by selecting the respective week.

Vanisa if you want to add something or explain it in more detail feel free Samples of collected tweets

Dependent variables

Hate speech: For our hate speech variables we collected data from Twitter. To collect tweets from the general users we utilize python codes 'GetOldTweets-python' from Jefferson Henrique which has been created in 2016.

The codes allow us to collect tweets that are older than seven days from the day we collect the data, which is a constraint we face if we use Twitter's official API. The python codes can be found below.

To make sure that the tweets are referring back to Trump statements, concerning the same topic we used the same keywords as we did for collecting Trump tweets.

The first variable we generated was the total number of tweets including at least one of the keywords. The collected tweets could either be positive or negative. This variable is supposed to represent how influential Trump is in defining relevant subjects of discussion and therefore guiding overall opinion formation. The more people talk about a topic the more controversial it seems to be.

To make sure that the tweets are referring back to Trump statements, concerning the same topic we used the same keywords as we did for collecting Trump tweets.

The first variable we generated was the total number of tweets including at least one of the keywords. The collected tweets could either be positive or negative. This variable is supposed to represent how influential Trump is in defining relevant subjects of discussion and therefore guiding overall opinion formation. The more people talk about a topic the more controversial it seems to be.

For our second variable we computed the average sentiments over all the collected tweets using the same keywords as Trump. For doing that we used a sentiment analysis

... Dung / Vanisa maybe you can add some more details - How many tweets did we collect per state ? ect ...

Hate Crime: The data on hate crime was provided to us by the FBI. The dataset includes all hate crimes that were reported as part of the Uniform Crime Reporting (UCR) program during the period of **???** to **???**. The data set describes the daily hate crime incidents within each state, including the type of the racial bias. Identical to the Trump tweets and hate speech we also focused on the number of weekly incidents within each state.

We would like to add that the number of reported incidents is likely to be lower than the actual number of hate crimes. In order to be included in the hate crime statistic, crimes had to undergo a two-stage decision process and had to meet a list of objective criteria, characterizing them as hate crime. In addition to this detailed process, incidents that would categorize as hate crime, might not have been reported to the police at all.

... Dung maybe you can fill in the details? - For which years did we gat the data? (1990-2020??), if you want you can add that the numbers are different from the yearly reports online ...

Control variables

Twitter Use:

For measuring the Twitter use within each state we rely on data from GESIS datorium. Upon request we were able to access geotagged Twitter posts from the U.S., collected in a 6 months period, both in 2014 and 2015. The files contained the aggregated number of hashtags used per day and state. One difficulty we were facing was the large volume of the data. The data was structured in zip-files per month containing several zip-files per day (Gesis datorium). To be able to calculate the total number of hashtags per state we had to join the txt-files first. For doing that we used ......

... Dung maybe you can describe a little how you added the different files together ect? Where did we run it because it was so big? ...

Election Results 2016

As a first control variable we collected data on the election outcome for the 2016 election within each state. We extracted the data from a Wikipedia table. We used the count of electoral votes to identify if the majority of the population voted for Trump or for Hillary Clinton. We recoded the values to a dummy variable with 1 reporting a higher voting share for the republicans and 0 reporting a majority for the democratic party.

Demographics:

For descriptive and controlling purposes we included a list of demographic variables, which were all collected from the US Census Bureau website (see table **XY**). The data was taken from four different online tables and had to be cleaned afterwards. We chose to use 5-year estimates instead of yearly data to have a more stable representation. We assumed that 5-year estimates are less vulnerable to yearly fluctuation and more likely to represent the general demographic structure in each state. Until now the latest year of reporting is 2018, our data therefore covers the years 2013 until 2018. Some of our other variables are taken from the year 2020, which is not represented in the statistics. We argue that by using averaged 5-year data the demographic structure is likely to remain constant and can be seen as a good estimate for 2020 as well.

Results

Data exploration

Analysis

Discussion

Limitation

1. For our data analysis we only used data since Trump officially was in office on 20th of January 2017. We decided on restricting the time frame because officially holding the position as president might have changed people's perception of Trump and thereby his influence. His status might have changed from being a well-known public figure to being an important politician who holds judicial power. However, people were already exposed to anti-hispanic rhetoric during his presidential campaign. Our starting point of data collection for tweets about Hispanics therefore isn't in line with the actual starting point of him using negative Hispanic rhetoric. For the negative rhetoric addressed towards Chinese and Corona on the other hand, we observe the initial reactions when phrases like "China Virus" and "Chinese Virus" have first been used. Therefore, the underlying processes might have been different at the different time points, since anti-mexican attitudes already had time to develop. The results for our two modes for Mexicans and Chinese are therefore probably not really comparable.

2. Another thing that might influence the comparability of the two minority groups is, that the population of Chinese were only in the center of attention for a couple of weeks due to other incidences like the murder of Georg Floyd and Trump's conflict between Twitter and Trump. Long term permanent change towards negative attitude which would then transform into hate crime towards Chinese is therefore unlikely.

3. One limitation of our study is the representativeness of the results. In total relatively few adult Americans are using Twitter, namely only 22 %. Out of them only one-in-five is following Trump's twitter account (2019b (n.d.)). We argued that tweets sent by Trump have a larger reach and also influence non-followers. Due to homophily preferences in network formation there might still be a bias towards more republicans seeing his posts. Our results concerning H1 and H2 are therefore likely to be overestimated for American twitter users overall.

Appendix

Keywords for tweets about Mexicans	Keywords for tweets about Chinese
Border	China virus
criminal	China
immigration	Chinese virus
drug	Coronavirus
gang	Covid

Keywords for tweets about Mexicans	Keywords for tweets about Chinese
Make America	Pandemic
Mexico	W.H.O
Wall	World Health Organization
Rapist	
Human Trafficking	

Appendix Table 1: Keyword table

Variables	Description	Source
Trump Tweets	Frequency of negative tweets per week · ·	Trump Twitter Archive Brown, Brendan. The Trump Twitter Archive. http://www.trumptwitterarchive.com/
Hate Speech	Number of tweets using the same keywords as Trump (can either be positive or negative) Average sentiment of these tweets	Twitter
Hate Crime – Hispanics	Number of weekly incidents (by State)	FBI

Control Variables	Description	Source
Election Results 2016	Dummy0 = Majority of votes for Democrats1 = Majority of votes for Republicans	Wikipedia 2016 United States presidential election, https://en.wikipedia.org/wiki/2016_United_States_presidential_election#Statistical_analysis
Total Population	Total number of citizens	US Census Bureau 2018: ACS 5-Year Estimates Data ProfilesTable: DP05
Hispanic Population	Number of Hispanics Population (by State)	US Census Bureau 2018: ACS 5-Year Estimates Data ProfilesTable: DP05
Chinese Population	Number of Chinese Population (by State)	US Census Bureau 2018: ACS 5-Year Estimates Data ProfilesTable: DP05
Age	Under 2525 to 4445 to 6465 and over	US Census Bureau 2018: ACS 5-Year Estimates Data ProfilesTable: DP05
Education	Less than High School degreeHigh School DegreeSome college or Associate's degree Bachelor's degreeGraduate or professional degree	US Census Bureau 2018: ACS 5-Year Estimates Data ProfilesTable: DP02
Household Income	Mean Household Income	US Census Bureau 2018: ACS 5-Year Estimates Data ProfilesTable: DP03
Unemployment Rate	Percentage of unemployed	US Census Bureau 2018: ACS 5-Year Estimates Data ProfilesTable: DP03

Control Variables	Description	Source
State Poverty Rate	Percentage of population below poverty rate	US Census Bureau 2018: ACS 5-Year Estimates Subject TablesTable: S1701
Twitter Use	Total number of hashtags (by State)	Gesis Datorium https://data.gesis.org/sharing/#!Detail/10.7802/1166

Appendix Table 2: Variables

References

2019a, Pew Research Center. n.d. "10 Facts About Americans and Twitter." https://www.pewresearch.org/fact-tank/2019/08/02/10-facts-about-americans-and-twitter/.

2019b, Pew Research Center. n.d. "About One-in-Five Adult Twitter Users in the U.s. Follow Trump." https://www.pewresearch.org/fact-tank/2019/07/15/about-one-in-five-adult-twitter-users-in-the-u-s-follow-trump/.

Asch, Solomon E. 1975. "Effects of Group Pressure Upon the Modification and Distortion of Judgments." Edited by H Guetzknow. *Groups, Leadership, and Men*, 177–90.

Álvarez-Benjumea, Amalia, and Fabian Winter. 2018. "Normative Change and Culture of Hate: An Experiment in Online Environments." *European Sociological Review* 34 (3): 223–37.

Bicchieri, Cristina, and Erte Xiao. 2009. "Do the Right Thing: But Only If Others Do so." *Journal of Behavioral Decision Making* 22 (2): 191–208.

Blumer, Herbert. 1958. "Race Prejudice as a Sense of Group Position." *Pacific Sociological Review* 1 (1): 3–7.

Bursztyn, Leonardo, Georgy Egorov, and Stefano Fiorin. 2017. "From Extreme to Mainstream: How Social Norms Unravel." National Bureau of Economic Research.

Chong, Dennis, and James N Druckman. 2007. "Framing Theory." *Annu. Rev. Polit. Sci.* 10: 103–26.

Cialdini, Robert B, and Noah J Goldstein. 2004. "Social Influence: Compliance and Conformity." *Annu. Rev. Psychol.* 55: 591–621.

Conger, Kate, and Mike Isaac. 2020. n.d. "Defying Trump, Twitter Doubles down on Labeling Tweets." https://www.nytimes.com/2020/05/28/technology/trump-twitter-fact-check.html.

Crandall, Christian S, Amy Eshleman, and Laurie O'brien. 2002. "Social Norms and the Expression and Suppression of Prejudice: The Struggle for Internalization." *Journal of Personality and Social Psychology* 82 (3): 359.

Gervais, Bryan T. 2013. "Incivility in Online Political Discourse and Anti-Deliberative Attitudes: An Experimental Analysis."

Karakostas, Alexandros, and Daniel John Zizzo. 2016. "Compliance and the Power of Authority." *Journal of Economic Behavior & Organization* 124: 67–80.

Krupka, Erin, and Roberto Weber. 2008. "Why Does Dictator Game Sharing Vary? Identifying Social Norms in the Laboratory." Working paper.

Rogan, Joe. 2019. "Joe Rogan Experience #1236 – Jack Dorsey" 1236.

Schaffner, BF. 2018. "Follow the Racist: The Consequences of Trump's Expressions of Prejudice for Mass Rhetoric. Semantic Scholar."

Schaffner, Brian F, and Patrick J Sellers. 2009. *Winning with Words: The Origins and Impact of Political Framing.* Routledge.

Titley, Gavan, Ellie Keen, and László Földi. 2014. "Starting Points for Combating Hate Speech Online." *Council of Europe, October 2014.*

Wahlström, Mattias, and Anton Törnberg. 2019. "Social Media Mechanisms for Right-Wing Political Violence in the 21st Century: Discursive Opportunities, Group Dynamics, and Co-Ordination." *Terrorism and Political Violence*, 1–22.

Zitek, Emily M, and Michelle R Hebl. 2007. "The Role of Social Norm Clarity in the Influenced Expression of Prejudice over Time." *Journal of Experimental Social Psychology* 43 (6): 867–76.