Honors Projects in Data Science   Senior Honors Projects

4-22-2022

Twitter's Role in an Increasingly Polarized Political Climate; A Look into the 2020 US Elections

Leanne Kendall

Twitter's Role in an Increasingly Polarized Political Climate; A Look Into the 2020 US Elections

BY Leanne Kendall

ADVISOR • Dr. Suhong Li
EDITORIAL REVIEWER • Dr. Ramesh Mohan

_____

Table of Contents

Twitter's Role in an Increasingly Polarized Political Climate
Honors Thesis for Leanne Kendall

<u>ABSTRACT</u>

Amidst politically strained times, one might wonder what has cause such an exaggerated gap between the views of democrats and republicans. For years, research has suggested the US's voting population is becoming increasingly politically polarized, with one of the causes being social media. This study's purpose is to understand more about the role that social media plays in the polarization of parties in the US.

The study is comprised of the analysis of over 3,000,000 tweets from 9/22/2020 through 11/10/2020 that mention or are written by senate and presidential candidates. Natural language processing, network graphing, and sentiment analyses were utilized to draw conclusions about the polarity of political discussion in the 2020 election. Based on the analysis, there are two distinct "echo chambers"—one republican and one democrat – that are very rarely interacting with one another. There are also spikes in sentiment around election day, suggesting polarity increases close to this day. Finally, popular negative topics gain rapid traction on twitter compared to more neutral topics, showing their infectious nature. These findings will contribute to the existing body of research surrounding the effect that social media has on the polarization of parties in the US.

INTRODUCTION

Social media plays a large role in American society. It only makes sense, then, that politicians and citizens willing to discuss politics online have developed their own social media niche, and have been using it increasingly to campaign, inform, and debate. Because of social media's increasing presence in US Politics, it is important that we know and understand the effects this may have. According to current research, the use of social media has been contributing to the growing political polarization problem in the United States. Although the ways in which this polarization is measured differ, this essentially means that the two dominant political parties in the United States are growing more and more fond of their group and more and more hateful towards the other. The visual below shows how much the gap between the views of democrats and republicans has widened from 1994 to 2015.

Figure 1 – Current Political Environment

The two charts below illustrate that the overlap between political values of Democrats and Republicans (the purple area) shrunk between 1994 and 2017, as the share of Americans with ideologically consistent values increased.

Source: "Political Polarization, 1994-2017." Pew Research Center, Washington, D.C. October 20, 2017, https://www.people-press.org/interactives/political-polarization-1994-2017/.
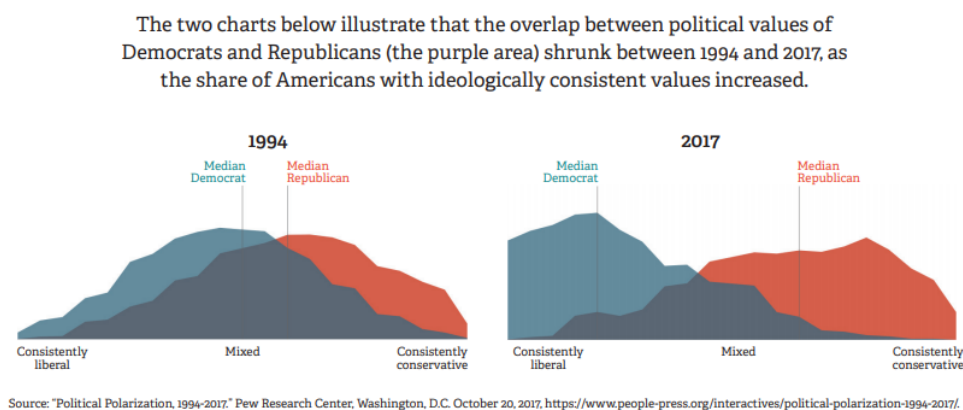
Figure 1 shows the gap between the political view of democrats and republicans has widened significantly from 1994 to 2017.

Growing polarization comes with increasingly severe consequences. Research has shown that political polarization has led to the following ramifications (Jilani 2019):

1. We're segregated in our own communities – residential communities are becoming more and more politically homogeneous.

2. Our political culture is more and more antagonistic – campaigns are more negative and focus more on tearing opponents down than building themselves up.

3. Our families are being undermined – families that cross political parties are seeing shorter dinners and less cross-partisan discourse.

4. We're losing trust in key institutions – higher education institutions, the press, the military, libraries, and other institutions are becoming less trusted.

5. Violence – Hate crimes have risen since the 2016 elections. We have also seen violent politically-driven crimes in recent years.

If the US were to continue down its current path, there may be more serious costs that jeopardize citizens' personal freedoms, their safety, and the stability of the nation as a whole. The data used in this analysis are tweets that mentioned a candidate for senate leading up to and right after election day in 2020. The study consists of a sentiment analysis, network analysis, and will examine the way in which polarized and non-polarized political information travels through the Twittersphere in order to gain a better understanding into why the US continues to become a more polarized nation.

LITERATURE REVIEW

There have been several articles published surrounding political polarization and social media. It is first important to understand the current political environment and One piece, titled "Polarization and Public Health: Partisan Differences in Social Distancing during the Coronavirus Pandemic", examines party differences that arose in response to the corona virus pandemic. It explains how prominent officials from either side of the political spectrum have sent divergent messages about the severity of the pandemic, with republicans such as Trump portraying the virus less severely than democrats to the public. After controlling for health, economic and other variables, their study found that people in counties with a higher percentage of republicans practice less social distancing. This study, then, suggests there is a behavioral gap due to the polarization of political parties in the US. It also suggests that messages sent from the media have an effect on political beliefs (Allcot, 2020). This study is relevant because it proves that political beliefs are (1) influenced by the media and (2) overlap with people's behavior in their daily lives.

Additional studies have shown that social media's role in political campaigns is expanding. One study by pew research center found that lawmakers posted far more content – and

received far more audience engagement – on social media during the 2020 election than in 2016 (Shah, 2021). This signifies that social media's influence in political campaign is increasing, making it something that is increasingly important to understand.

Figure 2 – Total Number of Interactions on Posts From Members of Congress During the 2016 and 2020 Elections
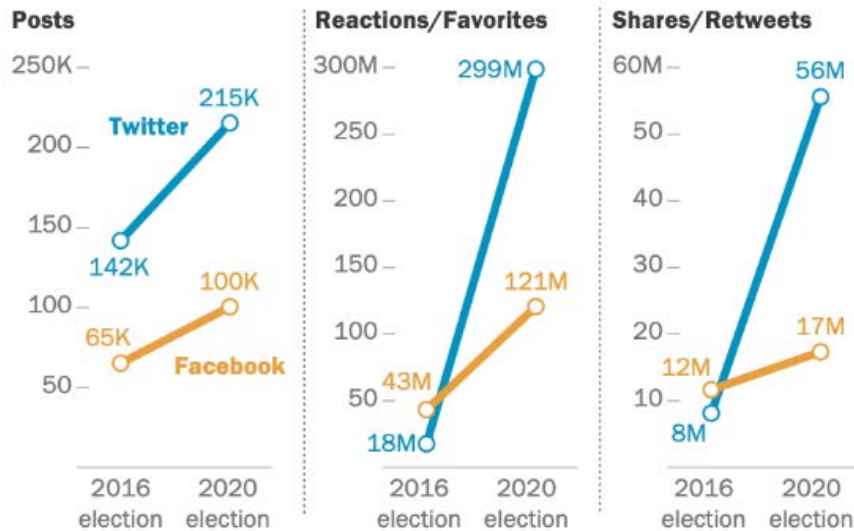

Figure 2 demonstrates the increased use of social media during elections from 2016 to 2020. There were more posts by members of Congress on both Twitter and Facebook, as well as an increase in reactions, favorites, shares, and retweets on both platforms. There were more 1561% more favorites and 700% more retweets on Twitter in the 2020 election.

Another insightful paper is "Measuring Online Political Dialogue: Does Polarization Trigger More Deliberation?". This article explores the relationship between online political deliberation and polarization using natural language processing of YouTube video comments. In order to measure the affective polarization of each comment, they conduct a sentiment analysis using the lexicon library ML-SentiCon, primarily for its reliability in the Spanish language. The article points to other researchers who have found that social media and the general internet environment can cause an increase in political or ideological polarization. This is due to what some authors have called "echo-chambers", where groups of like-minded

people form and avoid difficult or divergent discussions yet willingly discuss—and therefore strengthen—their similar views (Serrano-Contreras, 2020).

On social media, algorithms collect data about the user which, therefore, show users what they want to see –or things that they agree with. This creates a "group think" effect where beliefs will strengthen due to the inflow of agreed upon information. This occurs in relation to many topics, including politics. People often see news and ads that correspond to their current beliefs, which strengthens their political view and creates a more polarized political environment (Barbera, 2020).

One effect of this algorithm usage is the formation of online echo-chambers, which further strengthen political views. This is where groups of like-minded people form and avoid difficult or divergent discussions yet willingly discuss—and therefore strengthen—their similar views online.

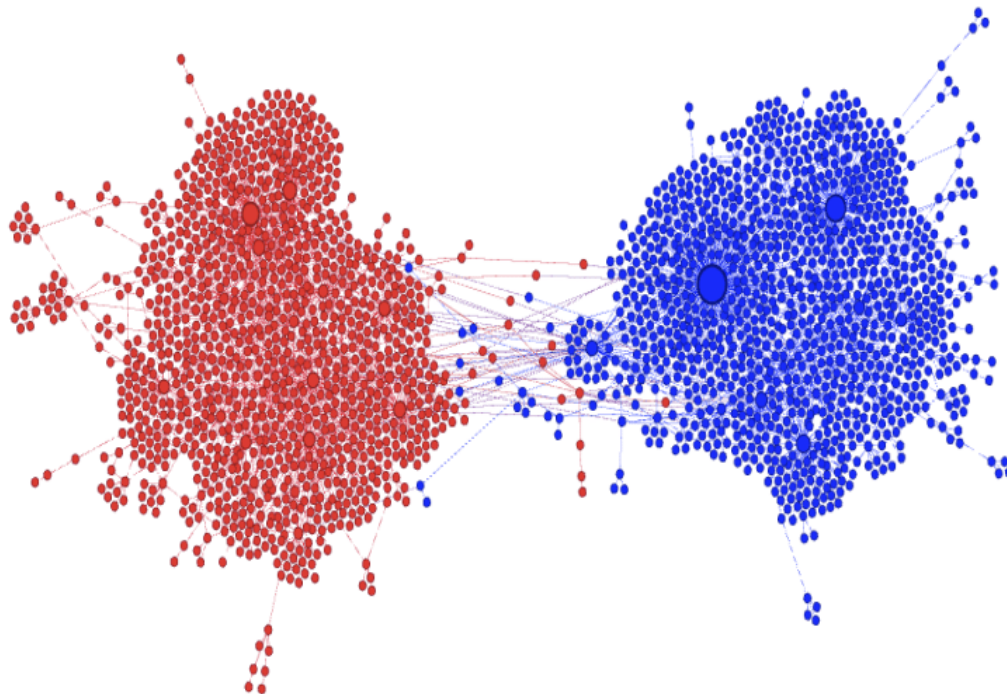*Figure 3 – Retweet Graph of Controversial Beef Ban Topic*

Figure 3 is a visual representation of the echo-chambers that form on social media.

Hence the name "echo-chambers", the graph shows how two communities engage with their own members and rarely engage with members of the other community. This is how opinions voiced on social media contribute to political polarization. Members of one community are surrounded by opinions that they agree with, resulting in a "group think" effect (Barbera, 2020).

Another article, "How Affective Polarization Shapes Americans' Political Beliefs: A Study of Response to the COVID-19 Pandemic", also studied the corona virus and its increasing role in American politics. Because the corona virus was originally a non-politicized entity, it was a good candidate to study how affective polarization actually translates into the differences in policy belief. Half of the respondents in the study were asked how they felt about the "United States'" response to the pandemic, while the other half were asked how they felt about "President Trump's" response. The study found that as affective polarization climbed for both democrats and republicans, there became less of a distinction between the United States responses and the Trump responses. This would indicate that as affective polarization increases, there is an increased inability to separate the United States (a seemingly superordinate category) from President Trump (Druckman 2020). This is relevant to my study in understanding the impact that affective polarization has in American politics. The finding that greater emotion can lead to the lack of distinction between the President and the country itself is an insight into affective polarization and its role in the US. This information contributes to the severity of the growing issue in the country.

A final relevant study was a Twitter sentiment analysis done on American partisan discussion of different COVID-19 related misinformation. The methodology used in this study is useful to me in learning more about how I could improve upon my methodology. For example, before conducting his sentiment analysis, the author deletes any twitter account with less than 50 followers or 50 friends and accounts that were younger than one month old in order to eliminate possible fake accounts. He also had to remove the word "Trump" as one of the words in the database, as it has a double meaning and, therefore, was skewing the sentiment positively (Havey, 2020). Though seemingly small, these things could have a huge effect on the results of my study, and should therefore be considered before beginning the analysis.

METHODOLOGY

The methodology for conducting this analysis is three-pronged: It consists of a network analysis (done in gephi), a sentiment analysis, and an investigation of how different topics gain traction or popularity on twitter. The first step in the entire process, however, is the data collection. This was done using a twitter listener that captured any tweet that mentioned or was written by a US Senate candidate from 9/22/2020 through 11/10/2020. The listener collected just over 3,000,000 tweets. They were stored in a JSON format and contained detailed information about each tweet, such as its time stamp, favorite and retweet count, user, user bio, user followers and following, user location, and more. The majority of the data analysis was done in python.

Network Analysis

The network analysis was conducted in gephi by connecting nodes through retweets.

Sentiment Analysis

The sentiment scores were attached to each tweet using the textblob library.

Topic Popularity

In order to decide on which topics to observe, I looked at the top ten hashtags in the dataset. Out of the top ten, three were polarized topics being used by one party to slander the other. I compared the trend in cumulative frequency by day by hour of these three topics against two neutral topics to see how the spread of non-polarized information compares to that of polarized information.
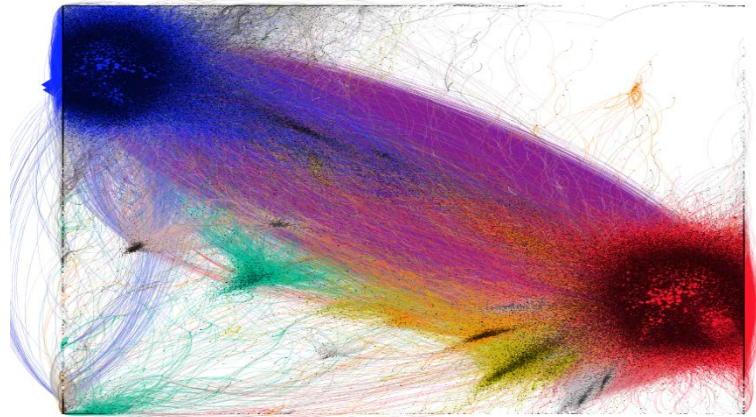
FINDINGS

In the same way the methodology was three-pronged, so are this study's findings. With each analysis different pieces of information were unveiled that contributed to the overall understanding of how political word on twitter is contributing to polarization.

Network Analysis

One node in the network graph represents a user. A line is drawn between nodes when one retweets the other. The retweet graph made in gephi shows two very defined communities: the republican and the democratic. It is clear that there is much more deliberation within these communities than there is from one community to the other. The graph essentially shows two echo chambers that were formed on twitter during the 2020 elections. Because of what we know based on our research, we can conclude that this echo chamber effect will cause people to view and interact with posts that reinforce their beliefs at a much higher rate than they will the opposing side. Therefore, their beliefs are bound to strengthen.

Sentiment Analysis

A sentiment score on a scale of -1 (very negative) to 1 (very positive) was attached to each tweet using the TextBlob library. To get a better understanding of the sentiment surrounding the tweets in the dataset I observed the average sentiment by day, the average sentiment by day and party, and the average sentiment by day and gender.
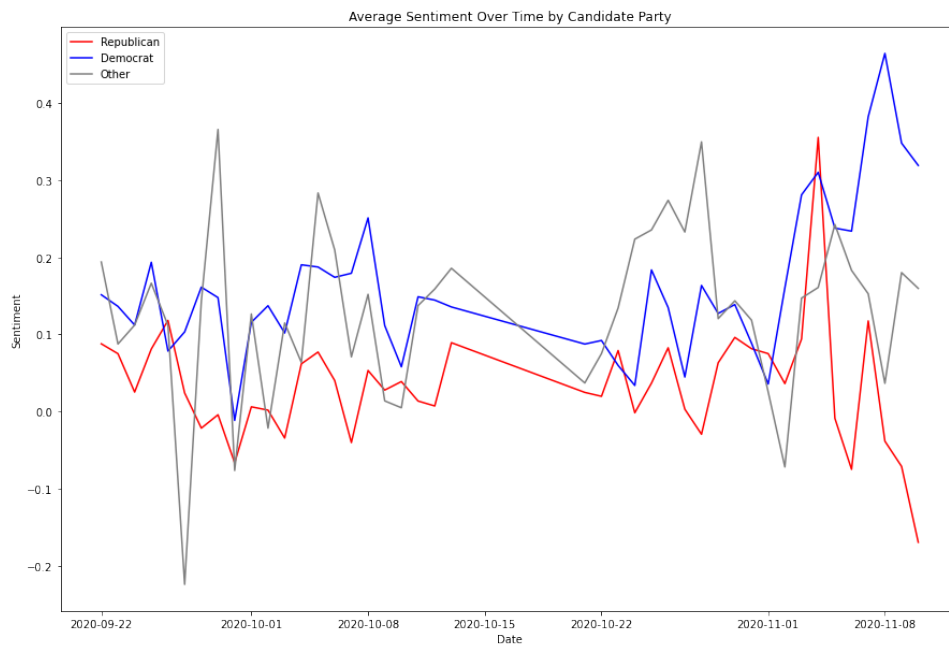
Figure 4 – Average Sentiment Over Time
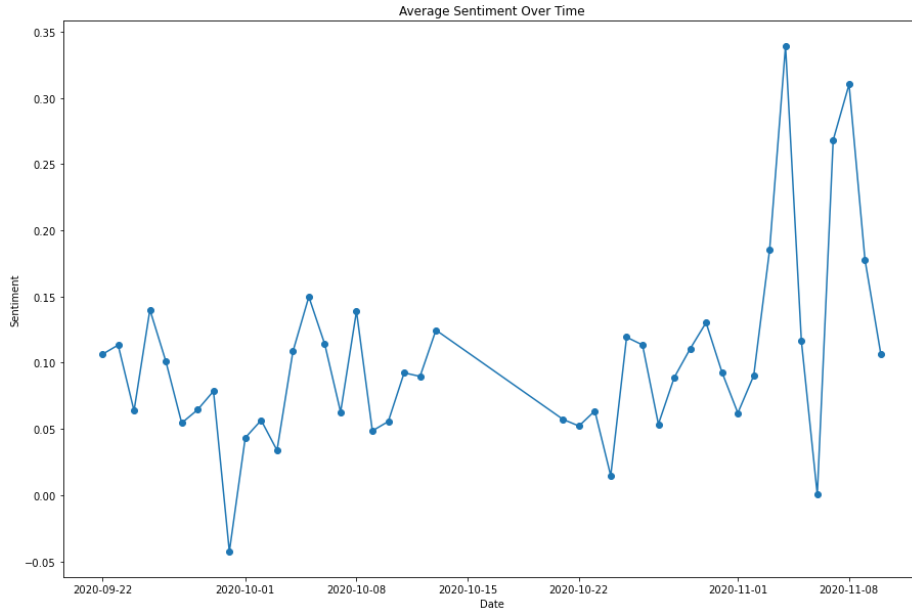
In figure 4, there are more extreme values (both positive and negative) near and after election day. This indicates that what is happening in politics outside of twitter is affecting the intensity of people's emotions.

Figure 5 – Average Sentiment Over Time by Party

Figure 5 shows the average sentiment by day and by candidate party (this means the graph indicates that the tweet mentioned a democrat or republican, respectively… and does not indicate the party of who the tweet was written by). The sentiment graph shows that the nation becomes more politically polarized right after the election, as we see the largest gap in sentiment on November 10. The graph also shows that in general, democrats were talked about more positively than republicans on twitter.

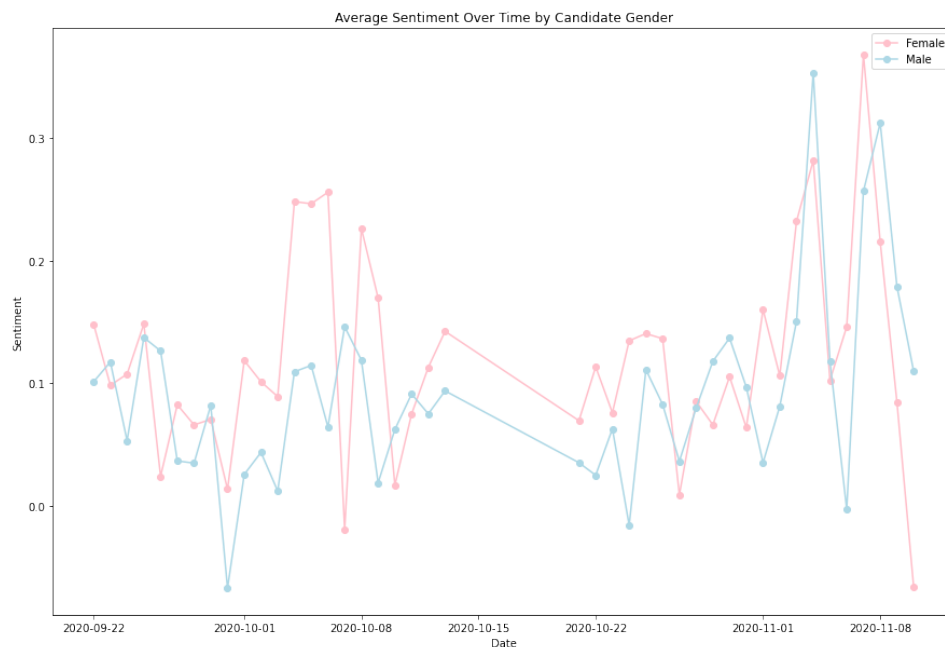*Figure 6 – Average Sentiment Over Time by Gender*

Figure 6 shows the average sentiment over time by the candidate gender. Again, this graph indicates the gender of the candidate being talked about and not that of the author of the tweet. There were no major discrepancies found in the sentiments, though it does appear that female candidates are talked about with a slightly higher sentiment than males.

Topic Popularity

The last piece of analysis sought to understand how topics gained popularity over time through this time period. The hashtags #vote and #election2020 were used as neutral baseline topics and the hashtags #LyingLindsey, #Expose2020, and #MoscowMitch were used as the polarized topics to look into. A comparison was then drawn between how the neutral topics gained popularity and how the polarized ones did by plotting each of their cumulative

frequencies by the hour. The reason I chose these topics was because they were amongst the top ten most frequently used in the dataset and, therefore, seemed most relevant to dissect.
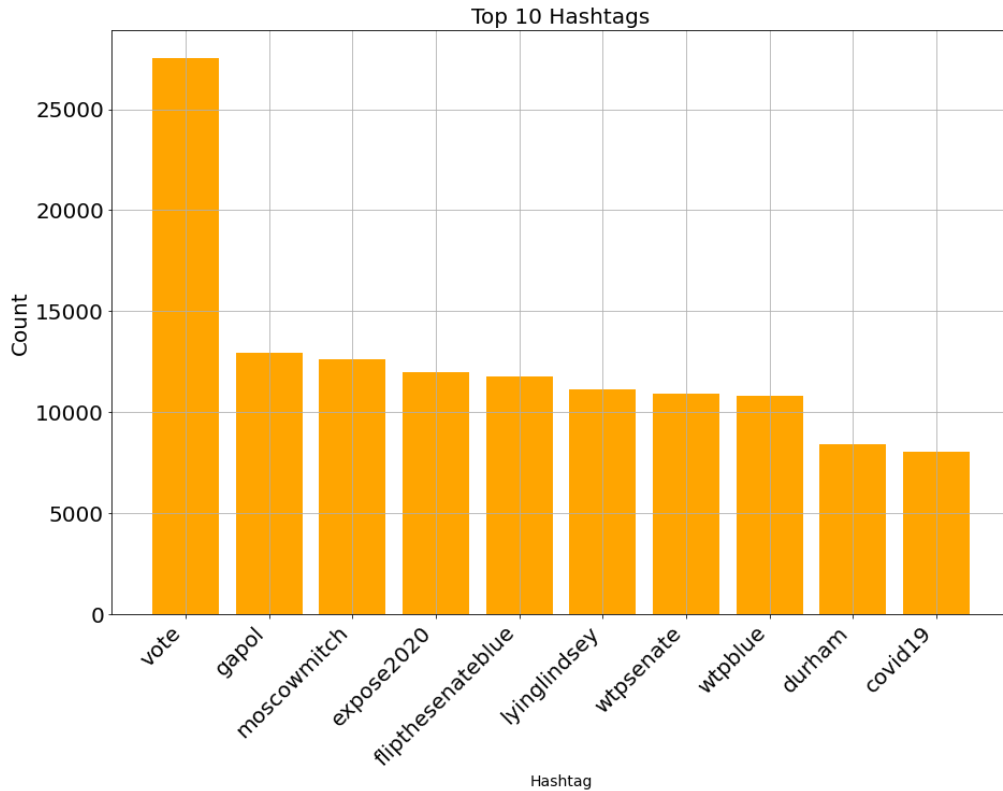
Figure 7 – Top 10 Most Frequent Hashtags

Figure 7 shows the top ten most frequent hashtags in the dataset. Vote, MoscowMitch, Expose2020, and LyingLindsey are amongst the top 10, and are all hashtags I decided to observe.

The two neutral hashtags, #vote and #election2020, show linear growth in their cumulative frequency while the polarized hashtags show more of a logistic growth that has a period of rapid spread and then a leveling off. This signals an element of "infectiousness" that polarized topics have and shows how the twitter platform is facilitating the more rapid spread of polarized information.
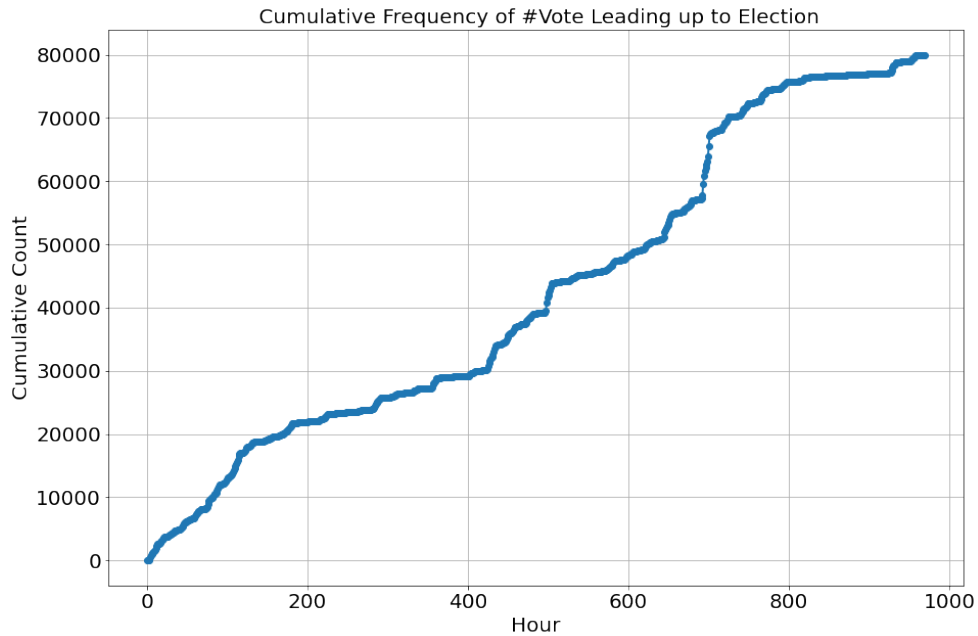
Figure 8 – Cumulative Frequency of #Vote Over Time

Figure 8 shows the frequency of #Vote through the time period. It displays a linear growth trend. This is one of two neutral topics I observed in order to get a baseline of how a non-polarized topic gains popularity on twitter.

Figure 9 – Cumulative Frequency of #Election2020 Over Time
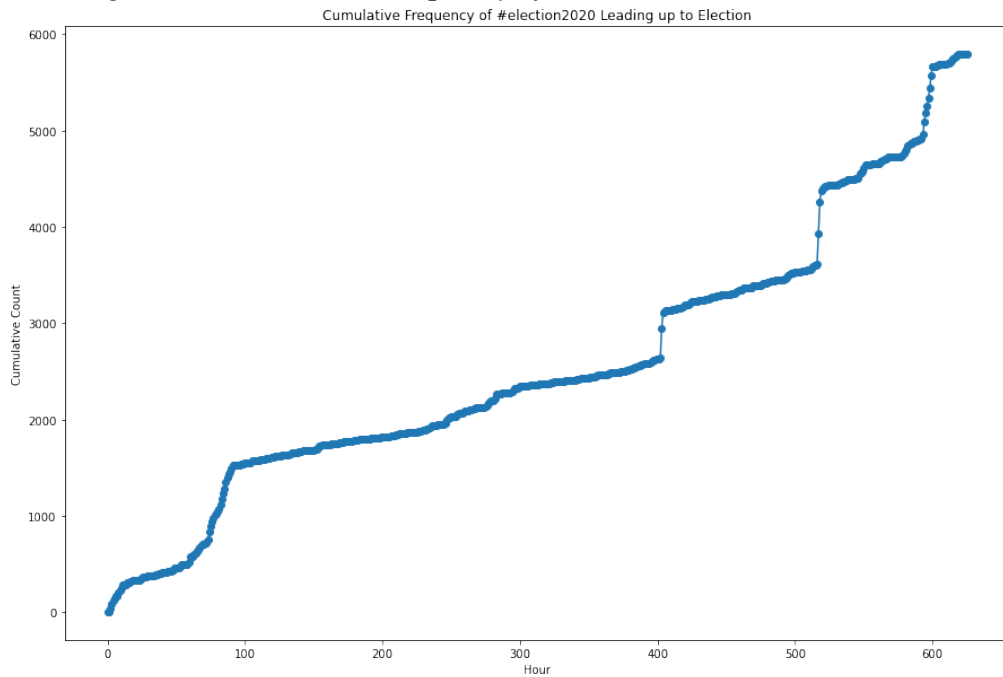

Figure 9 shows the hashtag #Election2020's cumulative frequency over time. It also follows a linear growth trend. I used this as a second neutral topic baseline.

The first polarized topic I observed was #LyingLindsey. The topic originated when Republican Senator Lindsey Graham changed his approval of the nomination of a Supreme Court Justice in the final year of Trump's term. Originally, he had said he would not support this nomination, then, upset democrats in September of 2020 when he changed his mind. This polarized topic shows a very different growth trend. It follows a logistic curve, where there is a period of rapid growth that then levels off over time.

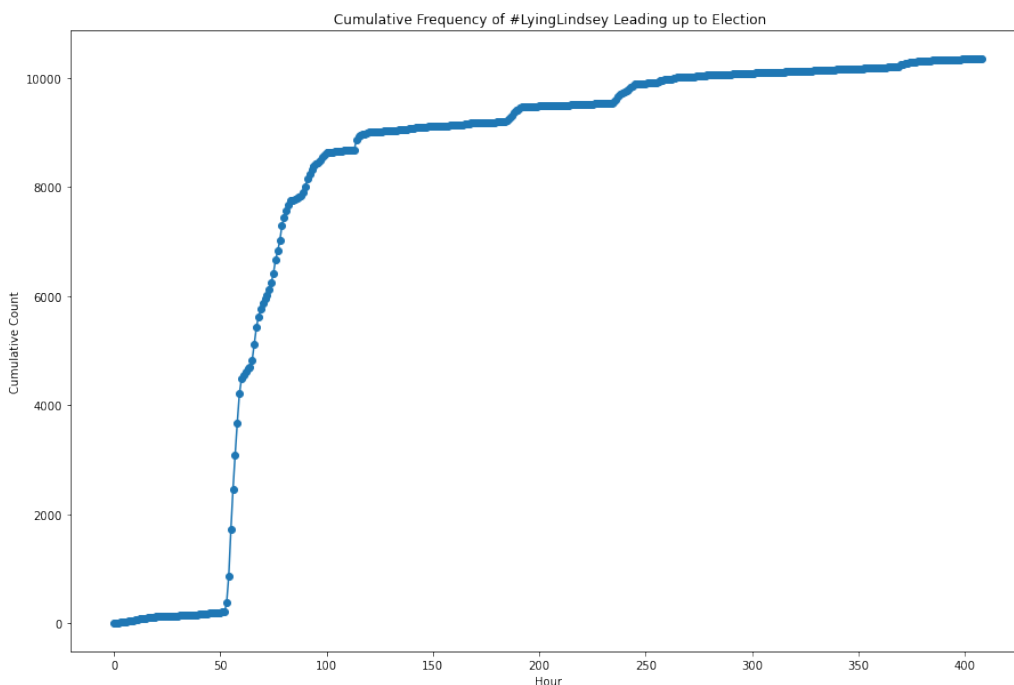*Figure 10 – Cumulative Frequency of #LyingLindsey Over Time*

Figure 10 shows rapid growth between 50 and 100 hours and then the growth levels off.

The next topic I observed was #Expose2020. This hashtag caught popularity when Project Veritas, a far-right activist group, who posts videos (often edited) that "expose" the government gained traction during the campaigns. Republicans used this hashtag to encourage the exposing of these government secrets. There is a similar logistic growth trend seen in this topic.

Figure 11 – Cumulative Frequency of #Expose2020 Over Time

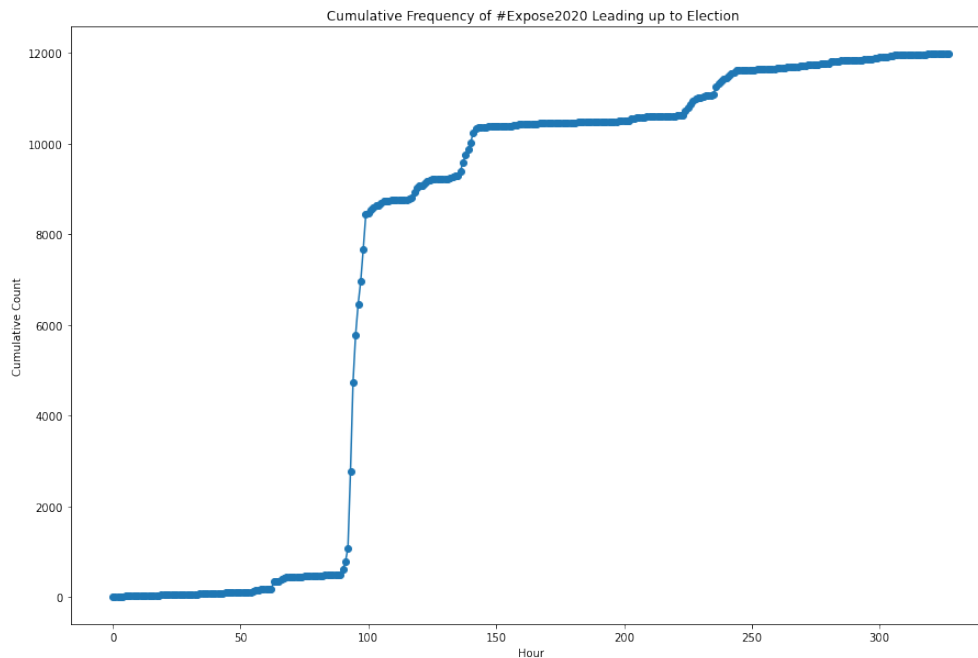Cumulative Frequency of #Expose2020 Leading up to Election

Figure 11 shows rapid growth between 90 and 105 hours and then the growth levels off.

The last polarized topic I observed was #MoscowMitch. This hashtag was used by democrats criticizing Mitch McConnell, minority leader of the senate, when he was accused of blocking stronger election security measures. Again, a similar pattern is seen in the way that this topic gains popularity. There is a period of rapid growth that levels off.

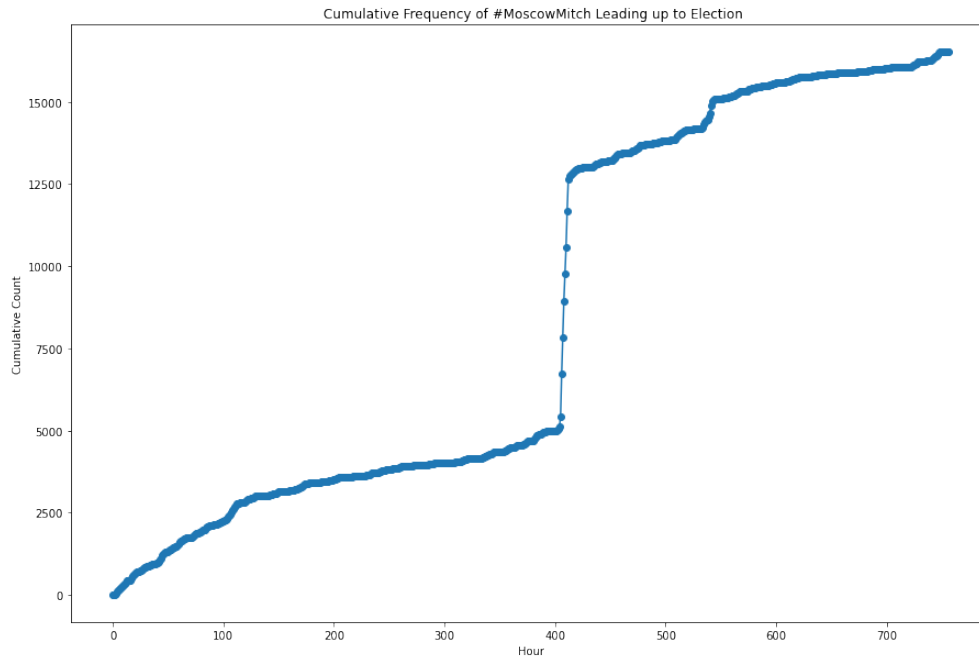*Figure 12 – Cumulative Frequency of #MoscowMitch Over Time*

Figure 12 shows rapid growth between 400 and 430 hours and then the growth levels off.

This analysis shows two different types of growth in popularity of neutral topics vs polarized topics. The two neutral topics grew steadily and followed a linear trend. The three polarized topics followed a logistic trend where there was a period of slow growth, followed by rapid growth, followed by slow growth. The disparity between the two types show that polarized topics are more "infectious" on Twitter than neutral ones. They gain traction rapidly rather than steadily. Twitter, then, is a platform where information that is polarized or controversial spreads faster than information that is not.

CONCLUSION

From the analysis, the following conclusions can be drawn:

Network analysis shows two echo chambers split by party affiliation on Twitter: The network graph produced results that showed much more in-group political discussion than inter-group discussion. Coupled with outside research that shows the effects of group-think and echo chambers on social media, the conclusion that these echo chambers lead to more polarized political views can be drawn.

As elections near, sentiment becomes more varied: Through the plot of average sentiment over the course of the election, it was seen that sentiment rises and falls sharply in the few days leading up to election day and the few days after. This indicates that the election itself impacts the degree of positivity and negativity that twitter-users have. This suggests that the election plays a significant role in the emotions of individuals who discuss politics on Twitter.

Polarized topics gain traction faster than neutral ones on Twitter: Through the analysis of how neutral topics gain popularity on twitter vs how polarized topics gain popularity on twitter, it was seen that polarized topics follow a logistic growth trend while neutral topics follow a linear one. This indicates a level of "insidiousness" for those topics that are more controversial, which leads to the conclusion that these Twitter is catered to spreading these types of topics at a faster rate than neutral ones.

These three findings all support the statement that political discussion on Twitter is contributing to the growing polarization in the US.

FUTURE RESEARCH

More research could be done to reaffirm the results that it produced. To add value to the results of this study, it would be interesting to look into fundraising efforts in political campaigns, and whether they are contributing to the flashes of activity that is seen in the polarized topics on Twitter. Additionally, identifying misinformation and evaluating how these tweets' popularity and sentiment compares to the average would give insight into how misinformation is spreading on Twitter.

<u>REFERENCES</u>

Allcot, H. (2020, April). Polarization and Public Health: Partisan differences in ... Retrieved January 22, 2022, from https://www.nber.org/system/files/working_papers/w26946/w26946.pdf

Barbera, P. (2020, August 24). Social Media, Echo Chambers, and political polarization (Chapter 3) - social media and democracy. Retrieved February 22, 2022, from https://www.cambridge.org/core/books/social-media-and-democracy/social-media-echo-chambers-and-political-polarization/333A5B4DE1B67EFF7876261118CCFE19

Druckman, J., Klar, S., Krupnikov, Y., Levendusky, M., & Ryan, J. (2020, August 24). How affective polarization shapes Americans' political beliefs: A study of response to the COVID-19 pandemic: Journal of Experimental Political Science. Retrieved January 22, 2022, from https://www.cambridge.org/core/journals/journal-of-experimental-political-science/article/how-affective-polarization-shapes-americans-political-beliefs-a-study-of-response-to-the-covid19-pandemic/B52D17EA47CCC436E8B1B3E008CA2A79

Havey, N. (2020, November 2). Partisan public health: How does political ideology influence support for covid-19 related misinformation? Retrieved February 22, 2022, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7604541/

Jilani, Z., & Smith, J. A. (2019, March 04). What is the true cost of polarization in America? Retrieved from https://greatergood.berkeley.edu/article/item/what_is_the_true_cost_of_polarization_in_america

Serrano-Contreras, I., García Marín, J., & Luengo, O. (2020, October). Measuring online political dialogue: Does Polarization Trigger More Deliberation? Retrieved January 22, 2022, from https://www.researchgate.net/publication/344540856_Measuring_Online_Political_Dialogue_Does_Polarization_Trigger_More_Deliberation

Shah, S., Widjaya, R., Smith, A., Rivero, G., & Chapekis, A. (2021, October 20). Charting congress on social media in the 2016 and 2020 elections. Retrieved January 12, 2022, from https://www.pewresearch.org/politics/2021/09/30/charting-congress-on-social-media-in-the-2016-and-2020-elections/

Spohr, D. (2017, August 23). Fake news and ideological polarization: Filter bubbles and selective exposure on social media. Retrieved January 22, 2022, from https://journals.sagepub.com/doi/full/10.1177/0266382117722446