**Can Hate, Sentiment, And Emotion Metrics Indicate an Increase in Political Polarization From the 2016 U.S. Presidential Election to the 2020 U.S. Presidential Election?**

Christian J. Farlin

University of Arizona

POL 424: Politics in the Digital Age

Professor Shmargad

**Table of Contents**

**Table of Figures**

## Table of Tables

## Introduction

### *Motivation for Research Topic & Why It Is Interesting*

The concept of political polarization has fascinated me ever since I was introduced to the topic in my A.P. U.S. Government and Politics high school course. I took the course in 2020, where the Democrat and Republican parties were vying for the presidency, the Senate majority, and the House of Representatives majority. I experienced firsthand how social media sites, such as Facebook and Instagram, served as media by which supporters of the respective political parties traded insults and threats and exuded undefined hate. Although partisan animosity and candidate slandering at the governmental and legacy media levels has existed since at least the 1796 Presidential Election, the rapid development of accessible, internet-connected devices alongside the development of social media sites has yielded forums where ordinary citizens ferociously attack one another at an intensity that exceeds the hate exuded to political candidates by other political candidates and legacy media organizations. This increasing divide between political parties—often referred to as political polarization—"is deeper than at any time since the American Civil War" (Flamino et al., 2023). The ease by which supporters of political parties can simultaneously attack each other and reinforce their own beliefs through "echo chambers" has accelerated the division between America's two major parties and their supporters.

Although the increasing political polarization in entities associated with political and news media organizations is significant and justifiably warrants further research, I narrowed my research to focus on analyzing the increase in affective polarization amongst the American voting populace because this topic is best analyzed through my toolkit of YouTube Data Tools, Gephi, and Google Colab. Because the "deepest divisions are found among the most interested, informed, and active [U.S.] citizens" (Gentzkow), I will analyze this voting subpopulation.

I aspire to complement the research set forth by (Flamino et al., 2023) by studying political polarization through hate speech and “negative” emotions in the 2016 and 2020 presidential elections. Yingying Chen and Yuping Wang’s analysis of incivility on YouTube in “Misleading Political Advertising FuelsIncivility Online: A Social Network Analysis of 2020 U.S. Presidential Election Campaign Video Comments on YouTube” will also be referenced. The primary difference between my research and the research of Yingying Chen and Yuping Wang is that I will analyze specific sentiment, emotion, and hate metrics rather than incivility. Ethan Xia, Han Yue, and Hongfu Liu’s “Tweet Sentiment Analysis of the 2020 U.S. Presidential Election” added to my interest. Since “sentiment analysis can be used as a low-cost and easy alternative to gather political opinion” (Xia et al., 2021) on Twitter, I purport that it is an essential metric that can similarly capture political opinion and political polarization on YouTube.

***Why It Is Important to Study***

I decided to analyze the 2016 and 2020 presidential elections specifically due to their divisive consequences on American politics. I purport that the 2016 presidential election, which “unfolded against a backdrop of intense partisan division and animosity” (Doherty et al., 2016), constitutes a key inflection point in the role of social media in the U.S. election process. Although Democratic Presidential candidate Barack Obama first utilized social media to distribute resources to and connect with voters in the 2008 presidential election, the ubiquity of smartphones alongside the increasingly popular use of social media applications such as Facebook, Twitter, and Instagram caused the 2016 presidential election to permeate ordinary life. The indivisibility between politics and personal life alongside the rapid procurement of social media political content attracted American citizens who were either "casual" voters or nonvoters. I define a "casual voter" as a voter who votes a straight ticket for his party in elections but does not immerse himself in political discourse before elections.

The 2020 election compounded the fact that "partisans' views of the opposing party are now more negative than at any point in nearly a quarter of a century" (Doherty et al., 2016) with increasing tension stemming from the claim from both Democrat candidate Joe Biden and Republican candidate Donald Trump that "the 2020 election [was] a fight for the 'soul' of America" (Dastagir, 2020). Interestingly, the 2020 presidential election exhibits a phenomenon called "negative voting", where people's votes are "more strongly driven by negative attitudes toward opposing parties and candidates than by positive attitudes toward one's preferred party and candidate" (Garzia et al., 2022). I purport that negative voting is a consequence of political polarization because voters believe that the "other" party is too extreme to hold office and would rather vote for an objectively worse candidate from their party so that the "other" party does not assume office. For example, many Democrats voted for Joe Biden in 2020 because they disliked Incumbent President Trump—and not because they liked Joe Biden.

Therefore, the 2016 and 2020 presidential elections constitute ideal elections to measure the increase in political polarization. My general hypothesis for this research is that an increase in the average negative sentiment, average negative emotions (anger, disgust, and fear), and average hate metrics from YouTube comments from the 2016 presidential election to the 2020 presidential election indicates an increase in political polarization.

Although the 2024 Presidential Election just took place, I purport that the recentness of the election yields less research to build upon than for the 2016 and 2020 Presidential Elections. Also, I assumed that more recent coverage of an event tends to be more biased and that bias decreases over time. I stipulated that it is much easier for researchers to objectively analyze events when the event is less salient.

### *Significance Statement (Summary of Findings)*

This paper details my efforts in determining if hate, sentiment, and emotion metrics can indicate an increase in political polarization from the 2016 U.S. Presidential Election to the 2020 U.S. Presidential Election. I utilized the "YouTube Data Tools: The Video Co-Commenting Network" module to collect videos and comment data for two Gephi networks. The first network is the 2016 U.S. Presidential Election Video Network, and the second network is the 2020 U.S. Presidential Election Video Network. Each network consists of videos collected from August 1st to November 21st of their respective election year. I used a "Frankenstein's Monster" notebook to calculate sentiment, hate, and emotion metrics for each network. After I computed metric scores for each video in each network in Google Colab, I appended the metrics to the Gephi networks for subsequent analysis in Gephi. My Gephi analysis is split into two groups per network: video statistic analysis, which concerns views, likes, and comments; and metric analysis, which concerns the disgust, negative sentiment (NEG), and "hateful" hate speech metrics. These three metrics were selected due to their relatively high scores in both networks.

Analysis of the 2016 U.S. Presidential Election Video Network revealed two video clusters. The "eye" cluster consists of densely connected videos with high numbers of views, likes, and comments. The "rings" cluster consists of rings of sparsely connected videos with little-to-no views, likes, and comments—outside of a minute group of outliers. Importantly, the lines that connect one node (video) to another are referred to as a "co-commenting edge". A co-commentor is a YouTube user who commented on two videos in the same network. I argue that the existence of such clearly defined clusters is caused by YouTube's Recommended Video Algorithm, which recommends related videos with high levels of views, likes, and comments.

Unlike the 2016 U.S. Presidential Election Video Network, the 2020 U.S. Presidential Election Video Network possesses three video clusters. Although the "eye" and "rings" clusters remain, they are less pronounced due to the emergence of a third "intermediate" cluster. I purport that the "intermediate" cluster signifies the decrease in either the importance of high view, like, and comment counts in the YouTube Recommended Video Algorithm process, or a decrease in the utilization of the recommended videos list by YouTube users.

After comparing the average metrics computed for each network, I concluded that I am unable to determine whether the metrics can indicate an increase in political polarization from the 2016 U.S. Presidential Election to the 2020 U.S. Presidential Election. The primary causes behind this inconclusive result are the small unrepresentative sample sizes, a significant percentage of zero-comment videos within each network, the disproportionately large influence of nodes with few comments in metric calculations, and the utilization of just the first 10 comments of each video to calculate each video's metric scores. I advise future researchers to address these research limitations before they conduct additional research.

## Literature Review

My research topic is a novel approach to a field of study populated by groups of well-researched political scientists and their findings. I discuss seven previous papers that are related to my search topic below.

First, Sobur Raji conducted sentiment and emotion analysis on Twitter tweets concerning the 2020 U.S. Presidential Election in "Sentiment Analysis and Emotion Classification Of The

2020 United State[s] Of America's Election Using Twitter" to determine if social media data can be used to predict future and live events. He utilized the Twitter API to collect 23,000 tweets. After he cleaned the data by removing URLs, duplicates, special characters, and punctuation, he utilized the Python TextBlob and Text2Emotion packages to conduct emotion classification and sentiment analysis. He deduced that tweets concerning Joe Biden have the highest number of positive tweets, whereas tweets concerning Donald Trump have the highest number of negative tweets. Through sentiment analysis, Mr. Raji deduced that the most prominent sentiment metric of both tweets with the name Joe Biden and tweets with the name Donald Trump is "neutral" (NEU). After removing the "neutral" sentiment option, he concluded that Joe Biden garners more positive replies (65% of tweets were positive) compared to Donald Trump (62% of tweets were positive). Lastly, Mr. Raji calculated the emotion metrics for tweets concerning each candidate, Mr. Raji deduced that the highest emotion score for tweets concerning Joe Biden was "surprise" and that the highest emotion score for tweets concerning Donald Trump was "anger" (Raji, 2023). Mr. Raji concluded that Joe Biden is the most preferred candidate based off these results. He also concluded that social media data can be used to accurately predict future and live events due to Joe Biden winning the 2020 U.S. Presidential Election.

Second, Ethan Xia, Han Yue, and Hongfu Lui conducted sentiment analysis on Twitter tweets concerning the 2020 U.S. Presidential Election in "Tweet Sentiment Analysis of the 2020 U.S. Presidential Election". They utilized the Twitter API to collect a sample of over 260,000 tweets related to the 2020 U.S. Presidential Election, processed feature extraction, then applied the Multi-Layer Perceptron classifier to classify the tweets with a positive or negative sentiment. After analyzing their results, the three researchers procured four conclusions. First, the candidates had a very close negative to positive sentiment ratio, which contradicted popular poll results (Xia et al., 2021). Second, negative sentiment is both more common and more prominent than positive sentiment in the social media domain. Third, conducting sentiment trend analysis on social media can lead to the detection of some key events. Fourth, sentiment analysis is a low-cost and simple alternative to gather political opinions.

Third, a contingent of 10 political scientists heled by James Flamino analyzed Twitter data to determine if political polarization increased between the 2016 and 2020 U.S. Presidential Elections in "Political Polarization of News Media and Influencers on Twitter in the 2016 and 2020 US Presidential Elections". The researchers continuously collected data using the Twitter search API with the names of the two presidential candidates in the 2016 and 2020 presidential elections respectively. The 2016 dataset contained 171 million tweets sent by 11 million users, and the tweets were posted between June 1st and November 8th (election day) in 2016. The 2020 dataset contained 702 million tweets sent by 20 million users, and the tweets were posted between June 1st and November 2nd (election day) in 2020. Their methods consisted of news media URL classification, influencer identification algorithms, influencer type classification, similarity network analysis, and latent ideology estimation. The researchers procured five conclusions. First, the number of influencers affiliated with media organizations declined by 10%, and they were mostly replaced by influencers affiliated with centre and right-leaning political organizations (Flamino et al., 2023). Second, professional media influencers overall shifted away from independent journalism and toward extreme bias right and fake news. Third, the division of influencers and users into opposing echo chambers increased from the 2016 election to the 2020 election. Fourth, the polarization of influencers and users correspondingly increased, which was discovered by analyzing latent ideologies of influencers and users (Flamino et al., 2023). Lastly, new influencers from 2020 possessed more polarized ideologies than the influencers who persisted from 2016. This study, introduced to me in my POL 424 class, was the inspiration for both this research paper and my honors presentation.

Fourth, researchers Ankur Agrawal and Tim Hamling conducted sentiment analysis on tweets concerning the 2016 U.S. Presidential Election in “Sentiment Analysis of Tweets to Gain Insights into the 2016 US Election”. They collected tweets and their associated metadata from Twitter and performed sentiment analysis on each tweet. The researchers created algorithms with the Java and Pearl programming languages and utilized them to access the tweets, clean them, perform the sentiment analysis, and aggregate the results. They collected a total of 6,854,213 tweets, where 2,810,051 tweets mentioned Hillary Clinton and 4,044,162 mentioned Donald Trump. After collecting the Twitter data and comparing it to the 2016 U.S. Presidential Election Electoral College results, the researchers deduced that Twitter sentiments corresponded with 66.7% of the actual outcome of the Electoral College. Additionally, “the overall sentiment of all tweets collected leaned more positively towards Donald Trump than it did for Hillary Clinton” (Agrawal et al., 2021). Interestingly, the researchers analyzed how different geographical locations affected each candidate’s popularity, determined the most prevalent issues discussed in tweets, and computed the ratio of a state’s population versus the number of tweets gathered for each state.

Fifth, a contingent of researchers led by Alexandra Siegal analyzed Twitter tweets to determine if online hate speech increased over Trump’s 2016 campaign and its aftermath and to what extent in “Trumping Hate on Twitter? Online Hate Speech in the 2016 U.S. Election Campaign and its Aftermath”. The researchers collected over 750 million election-related tweets and almost 400 tweets from a random sample of American Twitter users for their dataset. They utilized both “machine-learning-augmented dictionary-based methods and a novel classification approach leveraging data from Reddit communities associated with the alt-right movement" to analyze their substantial dataset (Siegal et al., 2021). The researchers concluded that there was no "persistent increase in hate speech or white nationalist language either over the course of the campaign or in the six months following Trump's election" (Siegal et al., 2021). Although key campaign events and policy announcements generated brief spikes in hateful language, they quickly dissipated. Overall, the researchers found "no empirical support for the proposition that Trump's divisive campaign or election increased hate speech on Twitter" (Siegal et al., 2021).

Sixth, researchers Rao Hamza Ali, Gabriela Pinto, Evelyn Laurie, and Erik J. Linstead conducted sentiment analysis of Twitter tweets related to the 2020 U.S. Presidential Election in "A Large-Scale Sentiment Analysis of Tweets Pertaining to the 2020 US Presidential Election". The researchers initially captured 12 million tweets that mentioned either Joe Biden or Donald Trump from October 31st, 2020, to November 9th, 2020, through the utilization of the "Tweepy" Python package. After applying three filtration steps—non-English language exclusion, creation date exclusion, and duplicate removal--researchers reduced the dataset to 7,609,756 tweets. They used the "Valence Aware Dictionary and sEntiment Reasoner (VADER)" sentiment analysis tool to conclude that Donald Trump had a sentiment score of 0.001 (neutral sentiment), whereas Joe Biden had a score of 0.097 (positive sentiment). Additionally, tweets that mentioned both candidates had an average score of 0.041 (Ali et al., 2022). Interestingly, both candidates received a similar sentiment leading up to the elections—even though Donald Trump was discussed significantly more. Sentiment around Joe Biden became increasingly positive at the start of Election Day. However, Donald Trump's sentiment score and the sentiment of tweets that mentioned both candidates remained within the neutral sentiment threshold (Ali et al., 2022). Additionally, the researchers revealed that deleted tweets that were posted after Election Day were more favorable to Joe Biden, whereas deleted tweets that were posted leading up to Election Day were more positive about President Trump. Lastly, the researchers uncovered a correlation between the age of a Twitter account and the number of positive tweets posted about Joe Biden: the older a Twitter account was, the more positive tweets it would post about Joe Biden (Ali et al., 2022).

Lastly, a contingent of researchers led by Loris Belcastro analyzed voter behavior on social media during the 2020 U.S. Presidential Election in "Analyzing voter behavior on social media during the 2020 US presidential election campaign". This report contains a host of conducted analyses, but my review will discuss their emotion analysis portion—which contains both sentiment metric analysis and emotion metric analysis. The researchers utilized a subset of a public repository that contains a real-time collection of 2020 US presidential election-related tweets from December 2019 to June 2021. However, the researchers' subset dataset consists of 160 million tweets published from September 1st to October 31st, 2020. 18 million re tweets, 110 million are retweets, and 32 million are replies posted by approximately 29 million users.

## Data, Methods, Findings, and Results

### *Selection of Query Terms*

I utilized the "YouTube Data Tools: The Video Co-Commenting Network" module to collect videos and comment data. I analyzed YouTube content for two reasons. First, YouTube allows researchers like me to collect data. Second, YouTube is a popular media by which ordinary U.S. voters interact with one another through videos and video comments. The search terms that I used for my research are "2016 U.S. Presidential Election" and "2020 U.S. Presidential Election" to capture data concerning the 2016 and 2020 U.S. Presidential Elections respectively. For a logical comparison to be made, the two entities being compared must be similar in size and time. Therefore, both search terms have 10 iterations, comments from 500 videos, and a restricted timespan from August 1st to December 1st. As a result, I acquired two .gdf files for each search term: a network of videos with comments and a network of channels. Although the network of channels may lead to interesting observations, I narrowed the scope of my analysis to solely concern the 2016 and 2020 U.S. Presidential election Video Networks. The researchers used the SentiStrength tool to reveal the "existing relationships between user polarity and the sentiment expressed in referring to the two presidential candidates" (Belcastro et al., 2022). After conducting sentiment analysis, they concluded that, on average, Trump's supporters produced tweets that are significantly more positive tweets than those produced by Biden's supporters. Additionally, Biden's supporters "devote[d] a significant number of negative tweets to their opponent" (Belcastro et al., 2022). After conducting emotion detection analysis, the researchers concluded that Trump's supporters expressed joy and confidence about Trump while expressing fear about Biden's election. Conversely, Biden's supporters show both trust and anticipation in "having Biden as future president of the USA" (Belcastro et al., 2022). Notably, the tweets of Biden's supported had a "more marked presence" of negative emotions concerning Trump (Belcastro et al., 2022). The three prominent negative emotions were anger, disgust, and sadness.

***Collecting Metrics***

I created a Frankenstein's Monster-esque notebook that contains metrics from two Python notebooks:

1. The "sen-hat" notebook, which contains two sets of metrics. First, the notebook contains positive, neutral, and negative sentiment metrics. Second, the notebook contains hateful, targeted, and aggressive hate metrics.
2. The "emo" notebook, which contains emotional metrics: joy, sadness, anger, surprise, disgust, and fear.

I purport that political polarization is best represented by negative sentiment, negative emotions, and hate. Table 1 contains the metrics that my Frankenstein's Monster-esque notebook calculated for each network.

| Notebook | Collected Metrics |
|---|---|
| Sen-Hat Notebook | Negative Sentiment (NEG)<br>Neutral Sentiment (NEU)<br>Positive Sentiment (POS)<br>Hateful<br>Targeted<br>Aggressive |
| Emo Notebook | Fear<br>Anger<br>Disgust |

*Table 1: Metric Collection Breakdown*

The Python notebook calculates the metric scores with Pysentimiento emotion, sentiment, and hate analyzers, which use the first 10 comments for each video. If the video has less than 10 comments, then the notebook collects the available comments. After I ran my notebook and calculated the comment metrics specified in Table 1, I appended them to my .gdf networks for the 2016 Presidential Election and the 2020 Presidential Election respectively.

### *Gephi Analysis Overview*

I analyzed my 2016 Presidential Election and 2020 Presidential Election networks with Gephi. Each network consists of nodes and edges, where each node represents one YouTube video. Edges, which connect nodes to each other, represent the number of co-commenters between two videos. A co-commenter is an individual who commented on multiple YouTube videos in the network. I will apply the subsequent steps to both the 2016 and 2020 Presidential Election networks so that I can compare the networks in similar states.

I first determined if there were any clusters of connected nodes within each network by applying the Fruchterman-Reingold layout and by partitioning the nodes by Modularity Class. I determined that two clearly defined node clusters exist in the 2016 U.S. Presidential Network. However, the 2020 U.S. Presidential Network blurs clear distinction that defined the 2016 node clusters through the introduction of a third node cluster type.

Next, I analyzed if there were outlier videos that have comments that are abnormally hateful, negative, angry, or exuded high levels of disgust or fear with respect to the average metric scores. I color-coded the nodes that have a higher sentiment, emotion, or hate score than their respective average metric score for the 2016 and 2020 Presidential Election networks. I enlarged these nodes and assigned them a unique color that makes these nodes stand out. I also deduced whether the nodes fit into one of the two clusters for the 2016 network or one of the three clusters for the 2020 network.

Lastly, I compared the 2016 and 2020 Presidential Election video networks to determine if political polarization increased through several methods. First, I compared the average metric scores for the 2016 and 2020 U.S. Presidential Video Networks specifically. If the average metric values in 2020 increased on average, then I purport that an increase in my "negativity" scores indicate political polarization because more YouTube comments are classified as negative—in sentiment, emotional content, or hate—under content discussing the 2020 Presidential Election than under content discussing the 2016 Presidential Election. Next, I compared the number of nodes with higher-than-average metric scores from the 2 networks. I purport that, if observed, a larger percentage of nodes with above-average metric scores in the 2020 Presidential Election network indicates that political polarization increased. The subsequent sections delve into these concepts in greater detail.

### *2016 U.S. Presidential Election Video Network Findings*

*Gephi Network Introduction & Google Colab Average Metric Analysis*

My 2016 U.S. Gephi Network is comprised of 452 nodes (videos) and 4,467 edges (comments). However, only 242 of the videos have at least 1 comment. Out of the remaining 210 videos, 25 videos had comments restricted, 7 videos were removed for irrelevancy, and 178 videos had 0 comments. This information is summarized in Table 2.

| 2016 US Presidential Election Video Information | | | | |
|---|---|---|---|---|
| Total Number of Videos | Videos With At Least 1 Comment | Videos With Comments Restricted | Number of Removed Irrelevant Videos | Videos With 0 Comments (Comments Enabled) |
| 452 | 242 | 25 | 7 | 178 |

*Table 2: 2016 US Presidential Election Video Information*

I removed the 210 videos because my Frankenstein's Monster program did not collect any metrics for them. Thus, the average metrics are calculated from the remaining 242 videos. Table 3 reveals the average metric values. I included the raw decimal metrics for future computations and the rounded percentage versions for simplicity.

| **2016 US Presidential Election Mean Metrics** | | |
|---|---|---|
| **Metric** | **Mean Value (Decimal)** | **Mean Value (%)** |
| Negative Sentiment (NEG) | 0.354429125 | 35% |
| Neutral Sentiment (NEU) | 0.350450416 | 35% |
| Positive Sentiment (POS) | 0.295120458 | 30% |
| Hateful | 0.119941682 | 12% |
| Targeted | 0.05472354 | 5% |
| Aggressive | 0.035639777 | 4% |
| Fear | 0.00881484 | 1% |
| Anger | 0.054206123 | 5% |
| Disgust | 0.224800887 | 22% |

*Table 3: 2016 US Presidential Election Mean Metrics*

Interestingly, the average sentiment scores are almost equal for the 2016 U.S. Presidential Election Network comments. Since approximately 35% of the comments exhibit negative sentiment, I decided to analyze it further. Additionally, only 21% of the comments exhibited one of the three forms of hate speech. The general "hateful" speech has the largest mean value of the three hate speech metrics, which is why I decided to analyze it in greater detail. Lastly, the average disgust value is approximately 4 times greater than the average anger value and approximately 22 times greater than the average fear value. Therefore, I decided to further analyze the disgust metric in greater detail. Table 4 reveals the 3 metrics that I chose to analyze in Gephi in further detail. One can deduce that a metric was chosen by observing the black star placed to the right of the mean value as a percentage column.

| 2016 US Presidential Election Mean Metrics | | |
|---|---|---|
| Metric | Mean Value (Decimal) | Mean Value (%) |
| Negative Sentiment (NEG) | 0.354429125 | 35% ★ |
| Neutral Sentiment (NEU) | 0.350450416 | 35% |
| Positive Sentiment (POS) | 0.295120458 | 30% |
| Hateful | 0.119941682 | 12% ★ |
| Targeted | 0.05472354 | 5% |
| Aggressive | 0.035639777 | 4% |
| Fear | 0.00881484 | 1% |
| Anger | 0.054206123 | 5% |
| Disgust | 0.224800887 | 22% ★ |

*Table 4: The 3 Selected Mean Metrics for the 2016 US Presidential Election Network*

The next section of this paper is comprised of three segments: Modularity Class Analysis, Video Statistic Analysis, and Collected Metric Analysis.

## *Gephi Network Analysis & Results of Findings*

### **Modularity Class Analysis & Findings**

Before I delved into the video statistics and collected metrics, I wanted to determine if the videos fit into identifiable clusters. Although there are multiple ways of partitioning networks into clusters, I prefer to use the "Partition by Modularity Class" option because it groups the video nodes into communities. A community has two characteristics: one, the nodes in the community have many edges (comments) connect them together; and two, there are more edges connecting nodes in the community then there are edges connecting nodes from the community to nodes in a different community. Thus, communities consist of nodes that are densely connected within themselves yet are barely connected to other communities. If I calculated a high modularity value, then my network has multiple clearly defined clusters that appear "dense" due to the high volume of edges connecting nodes to other nodes in the same community. For reference, the modularity class metric can have values that range from -1 to 1, where a value closer to 1 indicates well-defined community structures. However, my 2016 U.S. Presidential Election Network boasts a modularity value of 0.049, which indicates a weak community structure. However, the video network possesses two defining characteristics: a dense purple "eye" comprised of videos connected to other videos by co-commenters and a set of "rings" containing videos with few co-commenters. Figure 1 depicts the two-cluster 2016 U.S. Presidential Election Video Network below. To clarify, a co-commenter is an individual who commented on multiple YouTube videos in the same network.

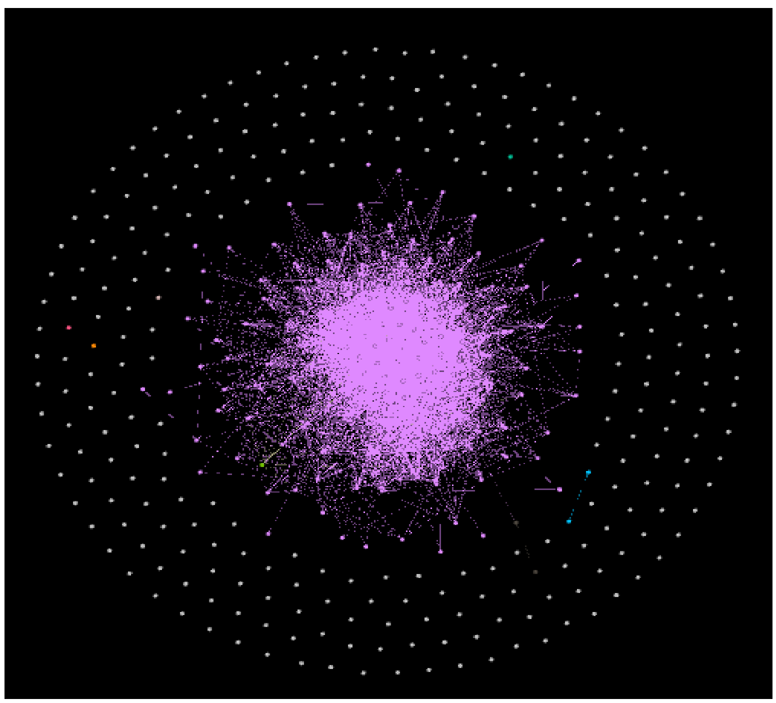

***Figure 1: 2016 Partition by Modularity Class Video Network***

The "eye" and "rings" clusters feature prominently throughout the video statistic and collected metric analysis because the two clusters share similar characteristics across the visuals. For the video statistic analysis section, the "eye" cluster video nodes possess higher views, likes, and comments than the "rings" clusters. In the collected metric analysis section, the "rings" cluster videos tend to have higher metric scores, yet the “eye” cluster videos tend to have a more consistent range of medium-to-low metric scores. Additionally, I purport that YouTube’s Recommended Video Algorithm created the densely connected “eye” network cluster. The subsequent sections analyze these phenomena and the possibility of a second hypothesis in greater detail.

**Video Statistic Analysis & Results**

The three measurable characteristics of a YouTube video are views, likes, and comments. These measurements are collected by “YouTube Data Tools: The Video Co-Commenting Network” module. Views are defined as the number of people who watched a video. Likes are defined as the number of people who clicked on the “like” button. Comments are defined as messages that people type into “comment” boxes and post underneath videos in the comments section. In Gephi, views, likes, and comments are represented by “viewcount”, “likecount”, and “commentcount” respectively. This section contains analysis of the number of comments, likes, and views that the YouTube videos in my 2016 U.S. Presidential Election Video Network.

<u>Comments</u>

The densely connected videos in the “eye” cluster have significantly more comments than the videos in the “rings”. As shown in Figure 2, most videos in the “rings” have fewer than 30 comments. Conversely, the videos in the “eye” have hundreds and even thousands of comments.

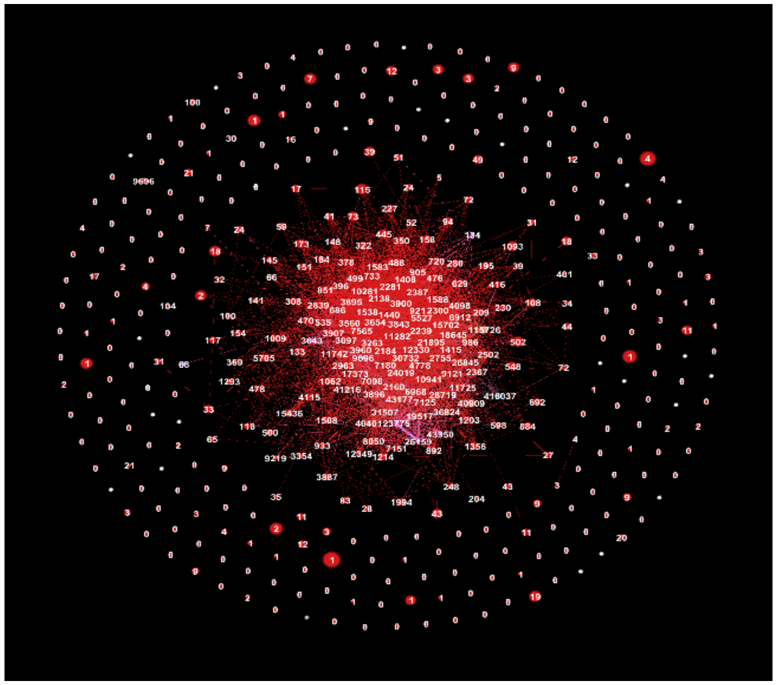

***Figure 2: 2016 Commentcount Video Network***

I purport that YouTube's Recommended Video Algorithm generates this phenomenon. After an individual watches a YouTube video, the YouTube Recommended Video Algorithm generates a list of recommended videos that are like what the individual watched in hope that the individual decides to watch one of the recommended videos instead of using the search bar to search for a different video. I purport that YouTube's Recommended Video Algorithm suggested that users who watched one "eye" video watch another "eye" video. Thus, the "rings" videos were recommended less. It can be inferred from Figure 2 that YouTube's Recommended Video Algorithm suggests videos with higher comment counts—even if a video with a lower comment count may be more like the initial video that the individual watched.

<u>Likes</u>

Similarly to the “viewcount” analysis, the videos in the “eye” cluster garnered significantly more likes than the videos in the “rings” cluster. Most videos in the “rings” have fewer than 100 likes. Conversely, the videos in the “eye” have hundreds and even thousands of likes. Figure 3 reveals this stark difference between the clusters.

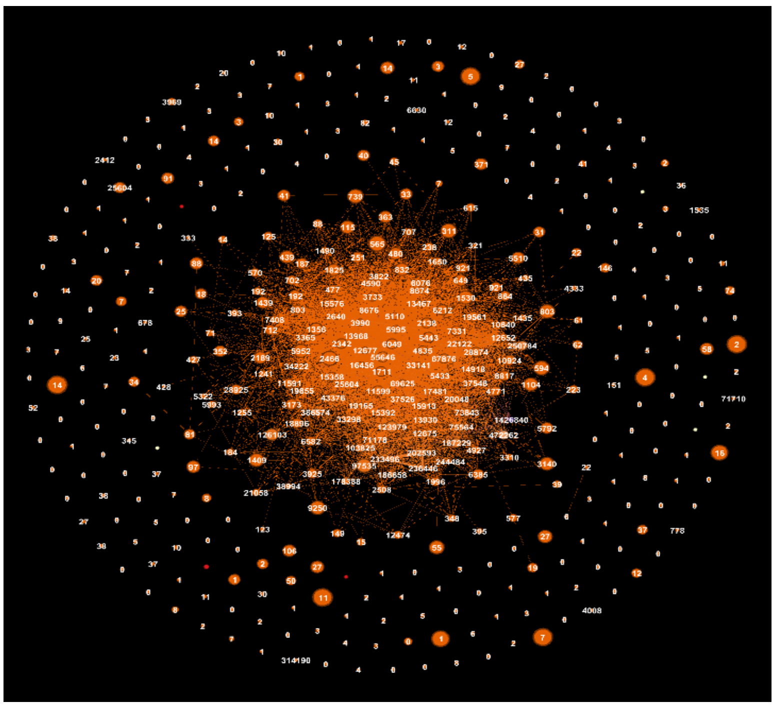

***Figure 3: 2016 Likecount Video Network***

I stipulate that the “likecount” discrepancy among the clusters indicates that YouTube’s Recommended Video Algorithm disproportionately favors videos with high like counts (Kapinos, 2024). In the context of this research, YouTube’s Recommended Video Algorithm tends to share “eye” cluster videos to individuals much more frequently than “rings” videos.

Views

The discrepancy between the “eye” and “rings” cluster is arguably the most pronounced in the “viewcount” analysis, which is revealed in Figure 4. Most videos in the “rings” have fewer than 800 views. Conversely, the videos in the “eye” have upwards of 106 million views.

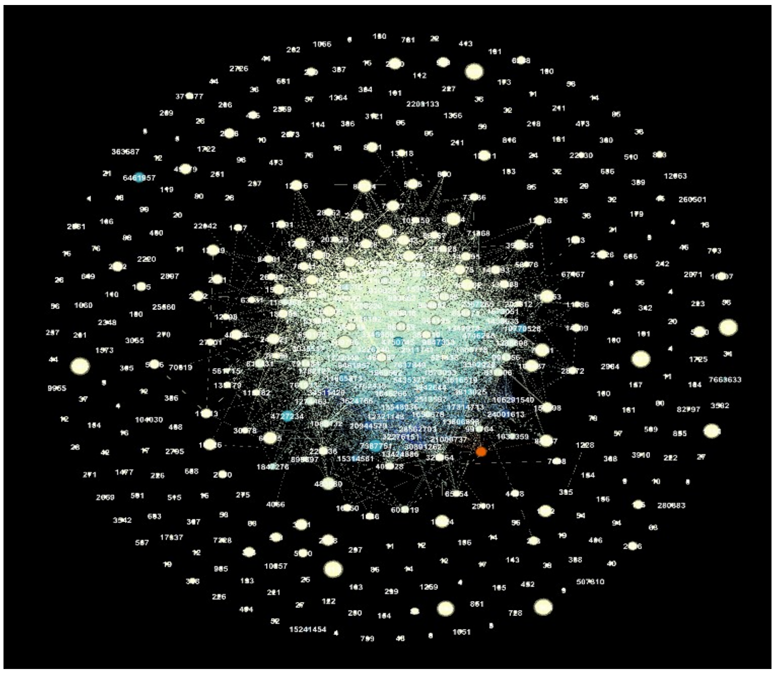

*Figure 4: 2016 Viewcount Video Network*

As with the “commentcount” and “likecount” analysis, the “eye” cluster videos accrue significantly more views than videos in the “rings” cluster due to YouTube’s Recommended Video Algorithm, which recommends videos with high view counts (Kapinos, 2024).

Video Statistic Analysis & Conclusion of Results

The above visualizations revealed that YouTube videos in the “eye” cluster possess much higher numbers of comments, likes, and views than videos in the “rings” cluster. The cause behind the starkly contrasting cluster video measurements is likely YouTube's Recommended Video Algorithm, which tends to recommend videos with high comment, like, and view counts. This consequently causes individuals who watched one video from the "eye" cluster and left a comment to watch another video from the "eye" cluster and to leave a comment, and this relationship is depicted in the visual as a line (edge) connecting two nodes. YouTube's Recommended Video Algorithm plays a much larger role in the videos that receive the most interaction with YouTube viewers than expected.

**Collected Metric Analysis & Results**

The three metrics that will be analyzed in this section are disgust, negative sentiment (NEG), and "hateful" hate speech. These metrics were selected due to their relatively high average metric values.

<u>Disgust</u>

Videos in the "rings" cluster tend to have higher disgust scores than videos in the "eye" cluster as depicted in Figure 5.

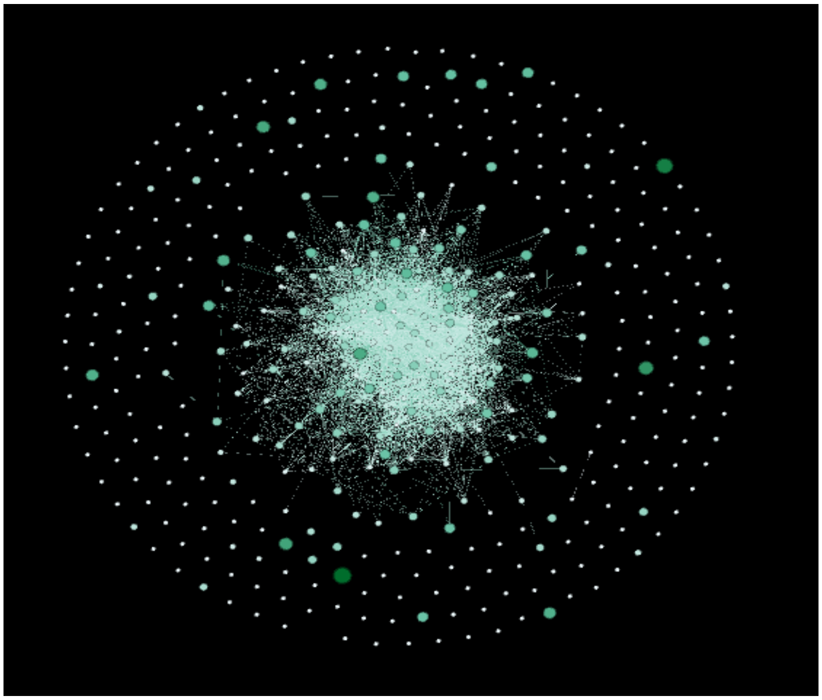

*Figure 5: 2016 Disgust Video Network*

However, as discussed in the commentcount analysis, videos in the "rings" cluster frequently possess fewer than 10 comments. Consequently, videos with fewer than 10 comments disproportionately affect the average disgust metric. Some of the videos with the highest disgust values, as indicated by the dark green node coloring, only have one comment. The fewer the comments, the more weight each comment has in the average metric calculation. Thus, videos in the "rings" have the highest variety in disgust values due to their small comment sizes (< 10). Conversely, videos in the "eye" have more accurate disgust metrics due to their large comment size (> 10).

Negative Sentiment (NEG)

Akin to my findings in the disgust metric analysis, videos in the "rings" tend to have higher "NEG" scores than videos in the "Eye" cluster due to the lack of comments that they possess. As shown in Figure 6, videos in the “rings” have the highest variety in disgust values due to their small comment sizes (< 10). Conversely, videos in the “eye” have more accurate “NEG” (Negative Sentiment) metrics due to the volume of comments available for metric calculations.

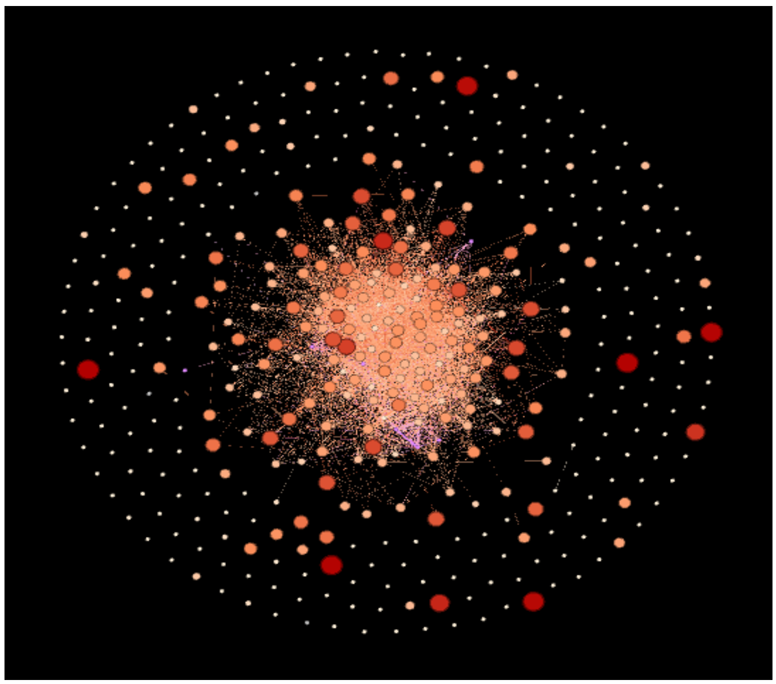

***Figure 6: 2016 Negative Sentiment (NEG) Video Network***

Although videos with fewer than 10 comments still disproportionately affect the average “NEG” metric, a significant percentage of “eye” videos have high “NEG” values. This discovery indicates that comment sections exude negative sentiment regardless of other metrics.

<u>Hateful</u>

Once again, videos in the “rings” tend to have higher “hateful” scores than videos in the “eye” cluster due to the difference in the number of comments. As shown in Figure 7, videos in the “rings” have the highest variety in “hateful” values due to their small comment sizes (< 10). Conversely, videos in the “eye” have more accurate “hateful” metrics due to the volume of comment available for metric calculations.

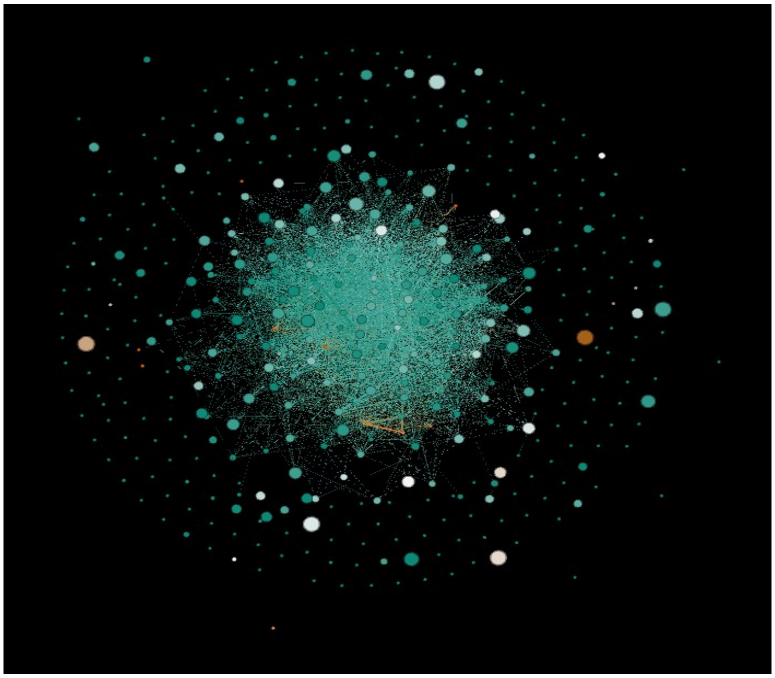

*Figure 7: 2016 “Hateful” Hate Speech Video Network*

Although videos with fewer than 10 comments still disproportionately affect the average “hateful” metric, a significant percentage of “eye” videos have medium-to-high “hateful” values. This discovery indicates that comment sections possess hateful comments regardless of other metrics.

Collected Metric Analysis & Conclusion of Results

The lack of comments on YouTube videos in the “rings” cluster has greatly skewed analysis of the network. I cannot reasonably state that videos in the “rings” cluster legitimately have higher metric values than those in the “eye” because the small handful of comments in the “rings” cluster videos disproportionately affect the metric score. For example, the YouTube video “2016 presidential election for economics awesome times 24/7” has only one comment. Consequently, the sole comment’s 0.96 negative sentiment value, or 96% negative, represents the comment section of the entire video. Thus, comparing videos with different comment counts is illogical due to the disproportionately high influence of comments in low-comment-count YouTube videos on metrics.

### *2020 U.S. Presidential Election Video Network Findings & Comparison*

#### *Gephi Network Introduction & Google Colab Average Metric Analysis*

My 2020 U.S. Presidential Election Video Network consists of 452 video nodes and 2,719 comment edges. Although I collected 452 videos, only 351 videos had at least one comment to generate metrics with. Of the remaining 101 videos, 51 had comments restricted and 50 had 0 comments. No videos were removed due to irrelevancy in this network. Table 5 summarizes these details.

| 2020 US Presidential Election Video Information | | | | |
|---|---|---|---|---|
| Total Number of Videos | Videos With At Least 1 Comment | Videos With Comments Restricted | Number of Removed Irrelevant Videos | Videos With 0 Comments (Comments Enabled) |
| 452 | 351 | 51 | 0 | 50 |

*Table 5: 2020 US Presidential Election Video Information*

I only included videos with at least 1 comment because I cannot make any statements on videos without comments. Therefore, my subsequent analysis concerns videos that have at least 1 comment and thus have calculated metrics. However, all 452 videos are present in my visuals to keep the shape of the network consistent. I collected a maximum of 10 comments per video just as I did for the 2016 U.S. Presidential Election network. Collecting more comments either crashed my computer or the Colab notebook. I removed the 101 videos because my

Frankenstein's Monster program did not collect any metrics for them. Thus, the average metrics are calculated from the remaining 242 videos. Table 6 reveals the average metric values. I included the raw decimal metrics for future computations and the rounded percentage versions for simplicity.

| 2020 US Presidential Election Mean Metrics | | |
|---|---|---|
| **Metric** | **Mean Value (Decimal)** | **Mean Value (%)** |
| Negative Sentiment (NEG) | 0.35107925 | 35% |
| Neutral Sentiment (NEU) | 0.372647516 | 37% |
| Positive Sentiment (POS) | 0.276273233 | 28% |
| Hateful | 0.069362688 | 7% |
| Targeted | 0.020591213 | 2% |
| Aggressive | 0.027475475 | 3% |
| Fear | 0.014301082 | 1% |
| Anger | 0.061337494 | 6% |
| Disgust | 0.200114368 | 20% |

*Table 6: 2020 US Presidential Election Mean Metrics*

Interestingly, the average sentiment scores are almost equal for the 20202 U.S. Presidential Election Network comments. Since approximately 35% of the comments exhibit negative sentiment, I decided to analyze it further. Additionally, only 12% of the comments exhibited one of the three forms of hate speech. The general "hateful" speech has the largest mean value of the three hate speech metrics, which is why I decided to analyze it in greater detail. Lastly, the average "disgust" value is approximately 3 times greater than the average "anger" value and approximately 20 times greater than the average "fear" value. Therefore, I decided to further analyze the disgust metric in greater detail. Table 7 reveals the three metrics that I chose to analyze in Gephi in further detail. One can deduce that a metric was chosen by observing the black star placed to the right of the mean value as a percentage column.

| 2020 US Presidential Election Mean Metrics | | |
|---|---|---|
| **Metric** | **Mean Value (Decimal)** | **Mean Value (%)** |
| Negative Sentiment (NEG) | 0.35107925 | 35% ★ |
| Neutral Sentiment (NEU) | 0.372647516 | 37% |
| Positive Sentiment (POS) | 0.276273233 | 28% |
| Hateful | 0.069362688 | 7% ★ |
| Targeted | 0.020591213 | 2% |
| Aggressive | 0.027475475 | 3% |
| Fear | 0.014301082 | 1% |
| Anger | 0.061337494 | 6% |
| Disgust | 0.200114368 | 20% ★ |

*Table 7: The 3 Selected Mean Metrics for the 2020 US Presidential Election Network*

The next section of this paper is comprised of three segments: Modularity Class Analysis, Video Statistic Analysis, and Collected Metric Analysis. Each section contains comparison analysis with respect to the 2016 U.S. Presidential Election Network

*Gephi Network Analysis & Results of Findings*

**Modularity Class Analysis & Findings**

After partitioning the 2020 U.S. Presidential Election Video Network by modularity class, a key distinction arises between the 2020 U.S. Presidential Election Video Network and the 2016 U.S. Presidential Election Video Network: the number of identifiable clusters. As mentioned above, the 2016 U.S. Presidential Election Video Network possesses two identifiable clusters: a dense "eye" in the center of the network, and a group of "rings" far away from the "eye". However, as shown in Figure 8, the 2020 U.S. Presidential Election Video Network possesses three identifiable clusters: a densely connected central "eye", a "middle" or "intermediate" cluster that bridges the "eye" and "rings" clusters, and the "rings" cluster.

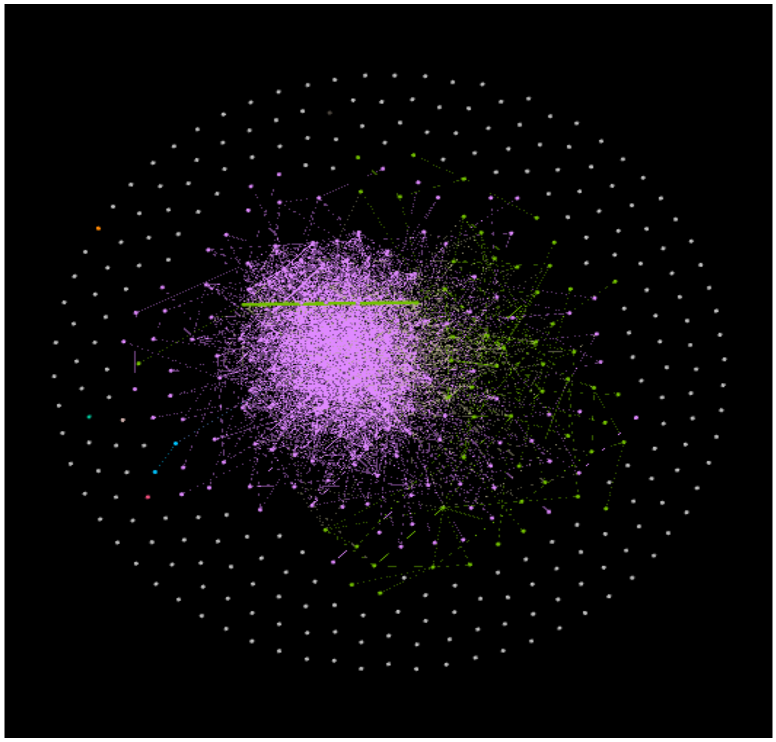

*Figure 8: 2020 Partition by Modularity Class Video Network*

Thus, the “eye” and “rings” clusters are less defined with respect to the 2016 U.S. Presidential Election Video Network due to the presence of the “intermediate” cluster. The “intermediate” cluster consists of a significant number of green nodes (videos) that “breach” the divide between the “eye” and “rings” with some co-commentors, but not as many as the less-defined “eye”. A key finding that will be explored in the video statistic analysis section is that I purport that the emergence of this third network characteristic indicates that YouTube’s Recommended Video Algorithm is less biased toward recommending videos with more likes, comments, and views.

## Video Statistic Analysis & Results

The three measurable characteristics of a YouTube video are views, likes, and comments. These measurements are collected by “YouTube Data Tools: The Video Co-Commenting Network” module. “Views” are defined as the number of people who watched a video. “Likes” are defined as the number of people who clicked on the "like" button. "Comments" are defined as messages that people type into "comment" boxes and post underneath videos in the comments section. In Gephi, views, likes, and comments are represented by "viewcount", "likecount", and "commentcount" respectively. This section contains analysis of the number of comments, likes, and views that the YouTube videos in my 2020 U.S. Presidential Election Video Network.

<u>Comments</u>

There are now three clusters of videos that possess different "commentcount" ranges as shown in Figure 9. First, the "ring" cluster is comprised of videos with fewer than 300 comments on average. Second, the "intermediate" cluster is comprised of videos that accrued between 300 and 3,000 comments on average. Lastly, the "eye" cluster is comprised of videos with more than 3,000 comments on average.

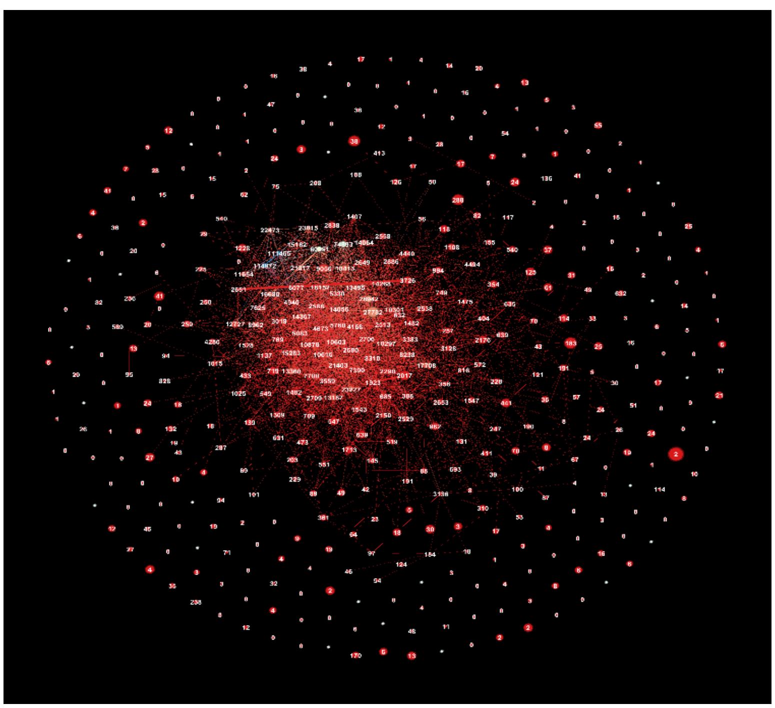

***Figure 9: 2020 Commentcount Video Network***

I purport that the emergence of the "intermediate" cluster suggests that YouTube's Recommended Video Algorithm has decreased the importance of the number of comments in the videos that it suggests to the YouTube user because individuals who watched an "eye" cluster video were suggested an "intermediate" cluster video. As reference, YouTube's Recommended Video Algorithm primarily relied on the number of comments in its video suggestion process in the 2016 U.S. Presidential Election Video Network, which consequently created the co-commenting edge-dense "eye" cluster.

<u>Likes</u>

There are no clearly defined classes based on like count as shown in Figure 10. This indicates that people relied less on YouTube's suggestions in 2020 than in 2016.

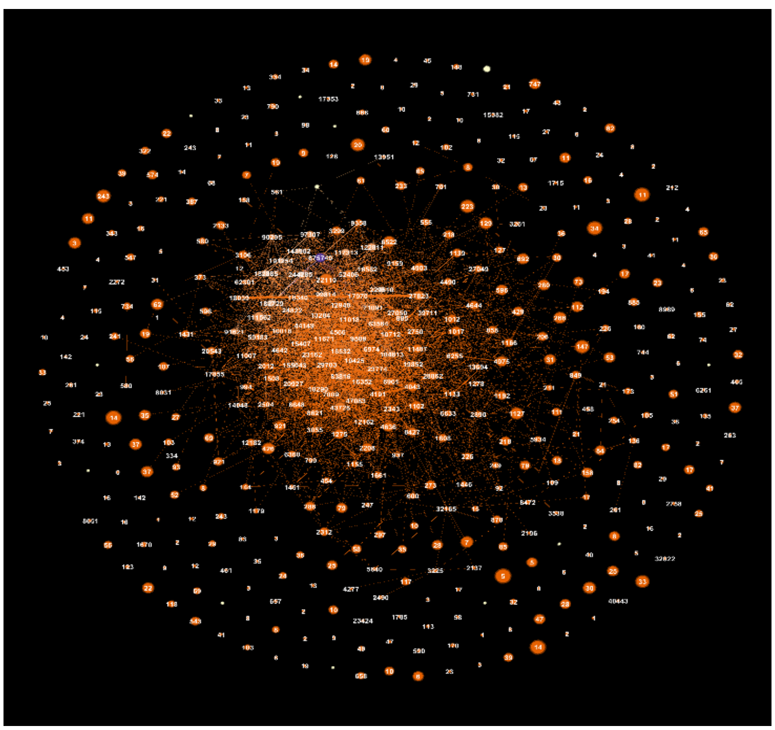

***Figure 10: 2020 Likecount Video Network***

The absence of clearly defined classes for like count indicates that videos with little-to-no co-commenting still experience relatively high like counts. The presence of "ring" outliers with tens of thousands of likes warrants further research in a future study. Although the "eye" cluster is present, it is overtaken by videos from the "intermediate" and "rings" clusters. Additionally, the lack of edges (connections) between videos in the "intermediate" and "rings" regions indicates that individuals may rely on their own searches for information--instead of merely clicking the next video suggested by YouTube.

<u>Views</u>

Unlike the 2016 U.S. Presidential Election Video Network, there are no clearly defined clusters based on view count as indicated by Figure 11.

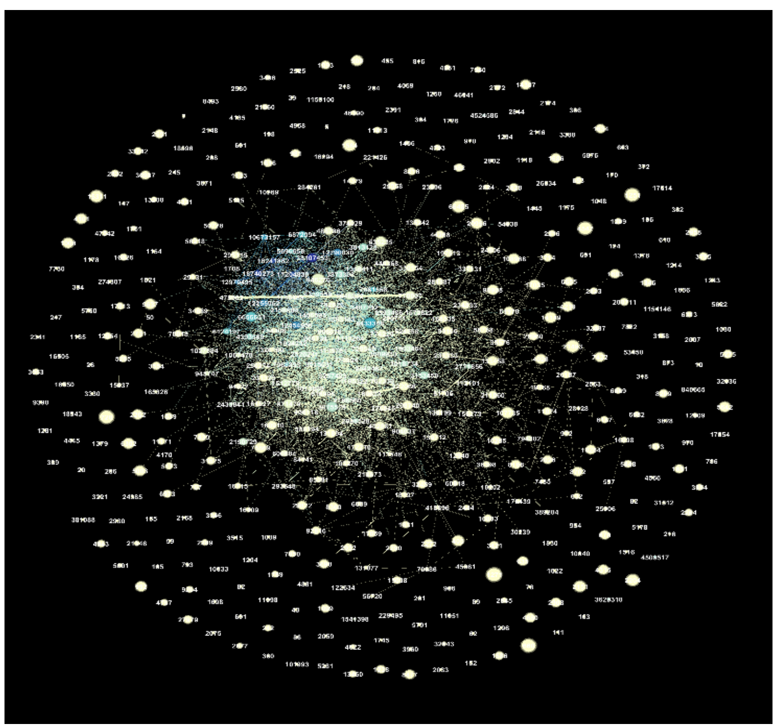

***Figure 11: 2020 Viewcount Video Network***

This indicates that some other factor affects which videos people watch. This undiscovered factor may be the extent to which a video has "gone viral" or the number of times the video was shared. The presence of "ring" outliers with trend-defying views and the unobservable factor both warrant further research in a future study.

Video Statistic Analysis & Conclusion of Results

Unlike the 2016 U.S. Presidential Election Video Network, YouTube's Recommended Video Algorithm appears to hold less influence over the videos that YouTube users select. Consequently, the introduction of the "intermediate" cluster blurred the distinctions between the once clearly defined "eye" and "rings" clusters. Thus, the "eye" and "rings" clusters introduced in the 2016 U.S. Presidential Election Video Network are less pronounced in the 2020 U.S. Presidential Election Video Network.

**Collected Metric Analysis & Results**

The three metrics that will be analyzed in this section are "disgust", "negative sentiment" (NEG), and "hateful" hate speech. These metrics were selected for two reasons. First, these three metrics were analyzed in the 2016 U.S. Presidential Election Video Network, and selecting the same three metrics to analyze in the 2020 U.S. Presidential Election Video Network allows me to compare and contract them. Second, these three metrics possess relatively high average metric values.

Disgust

Videos in the “rings” cluster tend to have higher disgust scores than videos in the “eye” cluster despite the emergence of the intermediate cluster as shown in Figure 12.

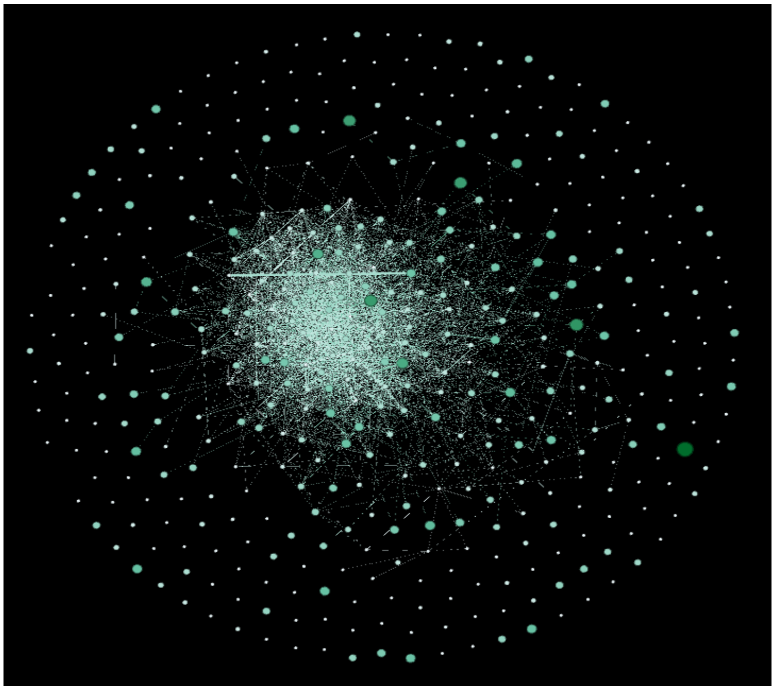

*Figure 12: 2020 Disgust Video Network*

Like the 2016 U.S. Presidential Election Video Network, videos in the “rings” cluster have the highest variety in disgust values due to their small comment sizes (< 10). Conversely, videos in the “eye” have more accurate disgust metrics due to the volume of comments available for metric calculations. Interestingly, the introduction of the third “intermediate” class does not significantly alter the distribution of the disgust values. Videos in the “eye” and “intermediate” classes still have smaller disgust values on average than videos in the “rings” class.

Negative Sentiment (NEG)

As shown in Figure 13, all three 2020 network clusters appear to have more videos with high “NEG” values than the videos in the two clusters in the 2016 U.S. Presidential Election video network.

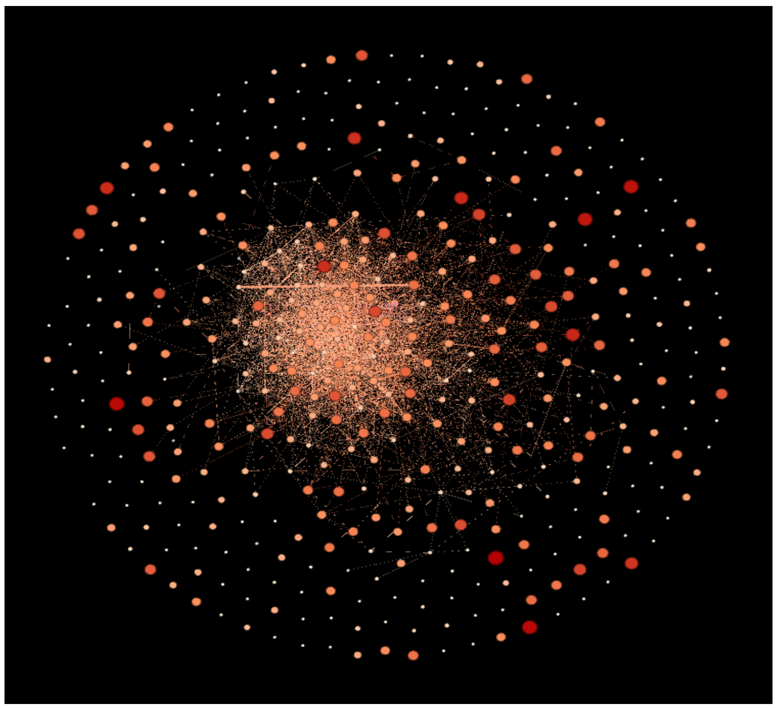

*Figure 13: 2020 Negative Sentiment (NEG) Video Network*

Interestingly, a significant number of videos in each of the three defined “classes” have higher “NEG” values than their 2016 counterparts. Additionally, unlike the disgust metric analysis, the “eye” cluster boasts videos that have high negative sentiment scores. Although the videos in the “rings” class possess the higher number of high-“NEG” videos, the overall average “NEG” score for the 2020 nodes with at least 1 comment is still approximately 35%--just like the 2016 U.S. Presidential Election Network.

Hateful

Videos in two of the three 2020 network clusters appear to have more high “hateful” values than in the 2016 U.S. Presidential Election video network. The 2020 U.S. Presidential Election Network Clusters and their “hateful” hate speech scores are displayed in Figure 14. The darker the shade of green, the higher the “hateful” score.

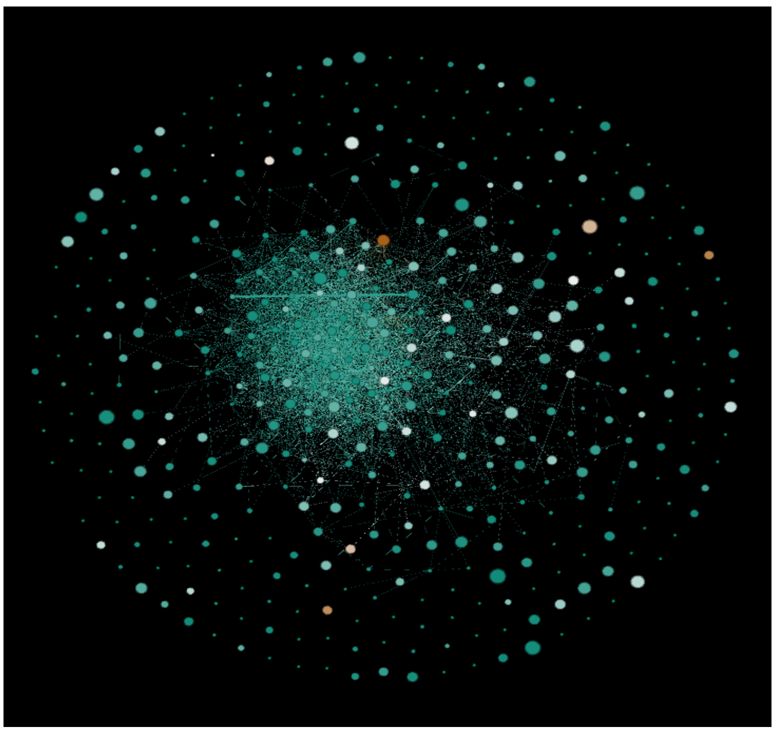

***Figure 14: 2020 “Hateful” Hate Speech Video Network***

The number of videos with high “hateful” videos in the “rings” and “intermediate” classes appears to have increased from the 2016 to the 2020 U.S. Presidential Election Video Network. Although the number of videos with high “hateful” values appears to have increased, the average hateful value for the 2020 U.S. Presidential Election Video Network is 5% less than it was in the

2016 U.S. Presidential Election. The 2016 average hateful value may have been disproportionately affected by outliers in the 2016 U.S. Presidential Election Video Network.

Collected Metric Analysis & Conclusion of Results

Similar to my conclusion in the 2016 U.S. Presidential Election Video Network collected metric analysis section, the lack of comments on YouTube videos in the “rings” cluster has greatly skewed analysis of the 2020 U.S. Presidential Election Video Network. I cannot reasonably state that videos in the “rings” cluster legitimately have higher metric values than those in the “eye” because the small handful of comments in the “rings” cluster videos disproportionately affect the metric score. For example, the YouTube video “Vote Counting Begins for 2020 U.S. Presidential Election” has only one comment. Consequently, the sole comment’s 0.947411537 negative sentiment value, or 97% NEG score, represents the comment section of the entire video. Thus, comparing videos with different comment counts is illogical due to the disproportionately high influence of comments in low-comment-count YouTube videos on metrics. Additionally, the introduction of the third cluster has eroded the clear distinctions between the “eye” and “rings” clusters.

**Results**

Table 8 reveals the percent change in the sentiment, hate speech, and emotion metrics that I analyzed.

| Percent Change In Metrics | | | |
|---|---|---|---|
| **Metric** | **2016 (%)** | **2020 (%)** | **Change (%)** |
| Negative Sentiment (NEG) | 35% | 35% | 0% |
| Neutral Sentiment (NEU) | 35% | 37% | 2% |
| Positive Sentiment (POS) | 30% | 28% | -2% |
| Hateful | 12% | 7% | -5% |
| Targeted | 5% | 2% | -3% |
| Aggressive | 4% | 3% | -1% |
| Fear | 1% | 1% | 0% |
| Anger | 5% | 6% | 1% |
| Disgust | 22% | 20% | -2% |

*Table 8: The Percentage Change in Sentiment, Hate Speech, and Emotion Metrics Across the Two Video Networks*

Surprisingly, negative sentiment held constant across the 2016 and 2020 U.S. Presidential Election Video Networks. As discussed in the introduction section, political polarization in the United States has created the largest ideological partisan divide since the American Civil War (Flamino et al., 2023). Additionally, political polarization reportedly increased in politically affiliated entities, news and other media organizations, and voters. Therefore, I expected to see an increase in negative sentiment. Similarly, I expected to see an increase in "hateful" hate speech and disgust. However, the hateful hate speech and disgust metrics decreased by 5% and 2% respectively, which indicates that political polarization decreased. The researchers who produced the studies in the literature review section primarily analyzed Twitter Data using Application Programming Interfaces (APIs), which serve as much better methods of accruing research data than the Co-Commenting module that I utilized. The discrepancy between the results of this study and the results of the studies discussed in the literature review stem from the limitations of this study, which are discussed in the conclusion.

After comparing the average metrics for the 2016 U.S. Presidential Election Video Network and the 2020 U.S. Presidential Election Video Network, I deduced that I cannot legitimately conclude that political polarization increased. The three metrics that I decided to analyze—disgust, NEG, and hateful—all decreased from 2016 to 2020 per the video networks. Additionally, the sample sizes of the networks are not representative of the population of YouTube videos concerning the 2016 and 2020 U.S. Presidential Elections specifically. Therefore, I could not conclude whether the hate, sentiment, and emotion metrics can indicate an increase in political polarization between the 2016 and 2020 U.S. Presidential Elections.

I indirectly deduced a second hypothesis concerning YouTube's Recommended Video Algorithm. The clearly defined "eye" and "rings" clusters observed in the 2016 U.S. Presidential Election Video Network is arguably a direct result of YouTube's Recommended Video Algorithm. The YouTube Recommended Video Algorithm is programmed to recommend a list of videos similar in content to the video that a YouTube user just watched. However, the algorithm is set to prioritize the recommendation of videos with high comment counts, view counts, and like counts. Importantly, the videos in the "eye" possess all 3 qualities that the algorithm recommends. Therefore, the algorithm is extremely likely to recommend videos in the "eye" cluster to a YouTube user who just watched a similar video. The algorithm will continue to suggest videos from the "eye" cluster if the YouTube user continues to watch videos with the same topic.

However, the introduction of the "intermediate" cluster in the 2020 U.S. Presidential Election Video Network indicates one of two things: a decrease in the importance of high comment counts, view counts, and like counts; or a decrease in the utilization of the algorithm's recommended video list by the YouTube user. First, it is likely that YouTube adjusted their Recommended Video Algorithm to recommend videos based on different features. For example, YouTube might have their algorithm prioritize the recommendation of similar-topic videos that “went viral” or were “trending”. Second, YouTube users in the 2020 dataset may just not click on videos in the recommended videos list. The first possibility indicates a significant change in how YouTube works, whereas the second possibility may just be an observation unique to these datasets. Regardless, my preliminary observations on this topic warrant future research.

## Conclusion

### *Summary of Findings*

I analyzed two YouTube Video Networks—one for the 2016 U.S. Presidential Election and one for the 2020 U.S. Presidential Election—to determine if sentiment, hate, and emotion metrics can detect an increase of political polarization amongst voters from the 2016 U.S. Presidential Election to the 2020 U.S. Presidential Election. I utilized the “YouTube Data Tools: The Video Co-Commenting Network” module to collect the videos and comments for each network. I crafted the networks in Gephi, then used Google Colab to calculate sentiment, hate speech, and emotion metrics for each video from the first 10 comments of each video in each network. Afterwards, I appended the Colab metrics to Gephi, then crafted the visuals included in this research paper. I utilized Excel to compute the average metric scores for each network. I decided to compare the video statistics—the number of comments, likes, and views—and selected Colab metrics—disgust, negative sentiment (NEG), and hateful hate speech (“hateful”)—for the two networks.

After analyzing the Gephi visualizations, I deduced that there were two node clusters in the 2016 U.S. Presidential Election Network: a dense “eye” comprised of videos with high numbers of comments, likes, and views; and a set of “rings” comprised of videos with low-to-nonexistent numbers of comments, likes, and views. YouTube’s Recommended Video

Algorithm is likely the culprit behind the formation of the “eye” cluster because it is programmed to show viewers a list of recommended videos. The list is comprised of videos with high numbers of comments, likes, and views, which are videos in the “eye” cluster.

Notably, videos in the “rings” cluster tended to possess larger metric values than videos in the “eye” cluster due to the lack of comments: the fewer the comments, the more influence each comment has on the video’s metrics. Thus, it is incorrect to state that videos in the “rings” have higher metric values on average because many “rings” metrics possess single-digit comments. The overinfluence of a single comment in the metrics for “rings” is evident in both the 2016 U.S. Presidential Election Network and the 2020 U.S. Presidential Network.

The 2020 U.S. Presidential Network possesses three node clusters: a dense “eye” comprised of videos with high numbers of comments, likes, and views; an “intermediate” green cluster that possessed middling numbers of comments, likes, and views; and a set of “rings” comprised of videos with low-to-nonexistent numbers of comments, likes, and views. The introduction of the “intermediate” column blurs the once-distinctive “eye” and “rings” clusters by serving as a “bridge” between them. The “intermediate” column reveals two key insights. First, YouTube’s Recommended Video Algorithm placed less emphasis on recommending videos with high levels of comments, likes, and views. Second, YouTube users rely less on the algorithm to select the next YouTube video to watch.

I was unable to conclude whether the stated sentiment, hate, and emotion metrics can detect an increase in political polarization from the 2016 U.S. Presidential Election to the 2020 U.S. Presidential Election due to three factors: the non-representative YouTube Video Networks; the overinfluence of nodes with few comments, and the inability to collect large quantities of YouTube videos. These factors are further explained in the limitations section. Thus, the revelation that the “disgust”, “negative sentiment”, and “hateful” hate speech metrics decreased between the 2016 and 2020 U.S. Presidential Elections are not statistically significant.

***Limitations of Work***

Five key limitations prevented me from reaching a conclusion. First, my samples are not representative of the population of YouTube videos concerning the 2016 and 2020 U.S. Presidential Elections. There are upwards of hundreds of thousands of YouTube videos concerning the 2016 and 2020 U.S. Presidential Elections, and my combined network video sample size consists of only 904 YouTube videos. This sample size shrinks to a mere 593 videos that Google Colab can calculate metrics for. Therefore, a key limitation of this study is the small, non-representative sample sizes for the networks.

Second, my computer was only able to process the computation of the first 10 comments of each video in my Google Colab metric calculations, which simultaneously caused videos with fewer than 10 comments to have a disproportionately high influence and videos with more than 10 comments to have a disproportionately low influence on the average metrics. Third, the metrics and average metrics only concerned the first 10 comments of each YouTube video. It is likely that the first 10 comments do not accurately represent YouTube video comment sections with hundreds of comments. Thus, the first 10 comments likely do not accurately represent videos with large comment section, which indicates that the calculated metrics for each video are not representative.

Fourth, the “YouTube Data Tools: The Video Co-Commenting Network” module is only able to capture a maximum of 500 videos and their associated comments at a time. To craft a representative sample, thousands of videos need to be captured at a time. This module does not contain a feature that prevents a user from obtaining duplicate videos, which renders running multiple queries with the same term inefficient. Fifth, the "YouTube Data Tools: The Video Co-Commenting Network" module did not let me restrict my query searches to videos with more than one comment. Consequently, my two video networks and their subsequent analysis included a significant number of videos with no metrics due to having zero comments.

### ***Direction for Future Research***

Future researchers who aspire to expand upon this topic must source a computing machine that can handle large volumes of comment data to ensure that the metrics that they collect are representative. Additionally, future researchers should consider utilizing a different module or video-collecting program. Python's "pytubefix" module may serve as a suitable alternative, as there does not appear to be a limit on the number of videos that it can collect. However, crafting a program that utilizes this module to collect the same attributes of each video as the Video-Co Commenting Network module requires an advanced knowledge of Python programming. Therefore, robust video-collecting software is required for future research on this topic.

Additionally, future researchers should consider locating and tracking a sample of YouTube channels across multiple election periods. Political polarization would be evident if the sentiment, hate, and emotion metrics that I utilized increased from one election to the next. Therefore, the analysis of both the Channel Network and the Video & Comments Network for multiple elections should yield more accurate results.

Lastly, future researchers should measure the change of these metrics across more than two presidential elections. Long-term trend analysis mitigates any temporary spikes or decreases in political polarization that may only be applicable to the time between two elections. By analyzing more elections, political polarization should be measured more accurately.

## References


Agrawal, A. & Hamling, T. (2021). Sentiment Analysis of Tweets to Gain Insights into the 2016 US Election. *Columbia Undergraduate Science Journal*, *11*(1). https://doi.org/10.52214/cusj.v11i.6359.

Ali, R. H., Pinto, G., Lawrie, E., & Linstead, E. J. (2022). A Large-Scale Sentiment Analysis of Tweets Pertaining to the 2020 US Presidential Election. *Journal of Big Data*, *9*(1), 79. https://doi.org/10.1186/s40537-022-00633-z.

Belcastro, L., Branda, F., Cantini, R., Marozzo, F., Talia, D., & Trunfio, P. (2022). Analyzing Voter Behavior on Social Media During the 2020 US Presidential Election Campaign. *Social Network Analysis and Mining*, *12*(1), 83. https://doi.org/10.1007/s13278-022-00913-9.

Chen, Y., & Wang, L. (2022). Misleading Political Advertising Fuels Incivility Online: A Social Network Analysis of 2020 U.S. Presidential Election Campaign Video Comments on YouTube. *Computers in Human Behavior*, 131, 107202. https://doi.org/10.1016/j.chb.2022.107202.

Dastagir, A. E. (2020, November 6). A Close Presidential Election Deepens the Nation's Divide. How do we Live Together Now? *USA TODAY*. https://www.usatoday.com/story/news/nation/2020/11/06/2020-election-american-divided-polarized-and-unsure-how-cope/6179404002/.

Doherty, C. (2016). Partisanship and Political Animosity in 2016. *Pew Research Center*. https://www.pewresearch.org/politics/wp-content/uploads/sites/4/2016/06/06-22-16-Partisanship-and-animosity-release.pdf.

Flamino, J., Galeazzi, A., Feldman, S., Macy, M. W., Cross, B., Zhou, Z., Serafino, M., Bovet, A., Makse, H. A., & Szymanski, B. K. (2023). Political Polarization of News Media and Influencers on Twitter in the 2016 and 2020 US Presidential Elections. *Nature Human Behaviour*, 7(1), 1–13. https://doi.org/10.1038/s41562-023-01550-8.

Garzia, D., & Ferreira da Silva, F. (2022). The Electoral Consequences of Affective Polarization? Negative Voting in the 2020 US Presidential Election. *American Politics Research*, 50(3), 303–311. https://doi.org/10.1177/1532673x221074633.

Gentzkow, M. (n.d.). Polarization in 2016. *Stanford University*. https://moodlearchive.epfl.ch/2022-2023/pluginfile.php/2714956/mod_resource/content/1/PolarizationIn2016.pdf.

Hayes, A. (2022, May 18). YouTube Stats: Everything You Need to Know In 2022! *Wyzowl*. https://www.wyzowl.com/youtube-stats/.

Raji, Sobur. (2023). Sentiment Analysis and Emotion Classification of the 2020 United States of America's Election Using Twitter. 10.13140/RG.2.2.36422.86087. https://www.researchgate.net/publication/375796633_SENTIMENT_ANALYSIS_AND_EMOTION_CLASSIFICATION_OF_THE_2020_UNITED_STATE_OF_AMERICA'S_ELECTION_USING_TWITTER.

Siegel, Alexandra A., Evgenii Nikitin, Pablo Barberá, Joanna Sterling, Bethany Pullen, Richard Bonneau, Jonathan Nagler, and Joshua A. Tucker. (2021) "Trumping Hate on Twitter? Online Hate Speech in the 2016 U.S. Election Campaign and Its Aftermath." *Quarterly Journal of Political Science, 16*(1), 71–104. https://doi.org/10.1561/100.00019045.

Xia, E., Yue, H., & Liu, H. (2021). Tweet Sentiment Analysis of the 2020 U.S. Presidential Election. *Companion Proceedings of the Web Conference 2021*. https://doi.org/10.1145/3442442.3452322.