Alex Toth
Dylan Herring
Shota Kajiya
Hayden Poore
Connor Saari
Capstone Paper
May 2nd, 2021
INFO 4700-001

Final Paper

Question:

How much misinformation was created or shared by Colorado's elected officials during the 2020 Presidential Election?

Motivation and Background:

The motivation behind this project was to develop a deeper understanding of the digital world that we are living in today. Social media plays a massive role in shaping public opinion around political candidates and public policies. Giving everyone a voice does not mean that everyone is going to use this voice in a responsible manner. Ahead of the 2020 presidential election social media platforms such as Facebook, Twitter, and YouTube promised to clamp down on misinformation shared on their platforms. The Capitol Riot on January 6th was perpetrated by supporters of Donald Trump who believed that he had won the election. Despite social media platforms labeling the Donald Trump election victory posts as misleading many people still believed it. This led to a riot at the United States Capitol Building in which over 140 were injured and 5 people lost their lives. This question is valuable to pursue because of the massive role that social media plays in our daily lives and understanding how to responsibly use these platforms. The methods that these platforms put in place ahead of the election to combat the spread of misinformation was clearly not effective enough. Our project takes a deep dive into understanding the flow of misinformation and how to better protect against it through in-depth interviews, surveys, and an analysis of election related tweets. In order to be able to fully understand the workings of the online political ecosystem we decided to focus on Colorado residents and Colorado representatives. By looking at a smaller sample size we were able to conduct more in depth research and better analysis by only focusing on our immediate environment. For the scope of this class a comprehensive research project looking at the entire nation would not have been able to yield us the extensive and complete analysis that our group completed over the course of the semester.

Ethical and Social Implications:

The ethical implications of our project encompass a few key points. We believe that the intentional spread of misinformation is unethical, and that people have a moral prerogative to stop the spread of misinformation whenever possible. The continual spread of misinformation that hurts people where they might otherwise be unharmed is a social, and societal issue. Misinformation is spread in a variety of ways, which we plan to address with this paper, but one pernicious means of spread is through social groups, both in person and online. People are the ones creating, sharing, and consuming misinformation. One unintended consequence of our project could be that people become even more distrustful of media sources, but still fail to learn ways in which they can recognize misinformation, so they become even more reliant on the media they already believe is truthful, even if it isn't.

Misinformation has deep-roots in the political sphere, which inherently allows for biases to be created. The findings that this project will uncover could possibly offer insights into how misinformation is created and what groups of people undertake it more often. One unintended consequence to these conclusions is that they will cast some as spreading misinformation, which is inherently an unfavorable characteristic. To better portray the findings from this project in a more agreeable manner, a standard of unbiased research and conclusions will need to be rigorously met.

Related Work:

There has been a flood of misinformation on social media sites in recent memory. It has become such a prevalent issue that the platforms feel the need to start addressing it. This has become an immense undertaking for these platforms. It has been a topic that has been researched thoroughly alongside the rise of these platforms. But because of the immense surge of misinformation related to the 2020 Presidential election, we are starting to learn more about how this type of information is spread. There is simply more misinformation to study.

The most effective way to study the amount of misinformation is to analyze the public posts containing misinformation. One method is to directly scrape the posts from the platform. These are posts created on public accounts on public platforms. It is fair game to use this information because it is publicly available. One way of doing this is using, "the Twitter Archiver add-on to search Twitter for tweets containing one or more of 11 common hashtags and three common key terms pertaining to the COVID-19 epidemic that were identified by the Symplur" (Kouzy 2). This project focuses on misinformation

pertaining to the COVID-19 pandemic. But it highlights how simple it is to filter posts. We will not pursue this method, rather we will be using the Twitter API. This can help us, for example, “obtain all links shared as part of a tweet that mentioned keywords related to the election (e.g. “hillary”, “clinton”, “donald”, “trump”, etc.) between October and November of 2016, representing a total of 24.1 million tweets” (Barberá 1). We will not be able to scrape and individually look at millions of Tweets, but this will be our primary method of data collection. We will delve further into this method of data collection in the Methods section below. Tweet scraping will be important because social media has become the most direct way politicians can communicate their beliefs, actions, and brand to the American people. If you want to know what a representative thinks of a political issue, there is most likely a tweet or Facebook post that conveys that sentiment. Because we are focusing on misinformation from Colorado Representatives, we will be able to pull their Tweets from their accounts and from a certain timestamp. This will be able to take us back into time and solely look at their activity during the peak of the 2020 election frenzy. These methods of tweet extraction give us real time data about CO representative’s posts that we can thoroughly fact check. This will help us directly answer our research question, as we are quantifying the amount of misinformation each representative from Colorado is projecting.

These two projects we have highlighted follow simple fact checking. There are many ways to fact check something that can be considered misinformation. Misinformation is defined as, “a “claim of fact that is currently false due to lack of scientific evidence”. It propagates without constraints, does not entail any curation or peer-review, and does not require any professional verifications” (Kouzy 2). Fact checking something means that there were witnesses and the fact can be verified by multiple sources. That means it happened. An example of fact checking can be seen in the COVID-19 study showing, “Tweets that contained genuine information regarding the COVID-19 epidemic were identified. Such information was cross-matched with the information presented by the World Health Organization (WHO), the Center for Disease Control and Prevention (CDC), peer-reviewed scientific journals, and prominent news outlets” (Kouzy 3). The information was verified through multiple agencies and stakeholders. Twitter identified these posts and regarded them as credible, being the truth. Regarding the 2020 presidential election, Twitter followed the same protocol. As Biden was projected the winner after one week of ballot counting, Twitter started identifying posts that conveyed that truth. The counter to these posts came, as with every political issue. But the flood of posts that refuted these claims became the subject of discussion. The polarization of our political system was on full display. We narrowed our research question to focus on this exact issue and to see how it may have been perpetuated in Colorado by our elected representatives and local constituents. Both of these studies take a look at both political parties, which is something we will have to as well.

These projects address which social groups are most susceptible to misinformation. This is important to note because we may see a distribution based on the data we have. But previous studies about election misinformation in 2016 demonstrate that, “age and partisanship were the two most predictive factors. Individuals of ages 65 and higher were nearly five times more likely to share false news stories on Twitter than those ages 18-25. Registered Republicans users were three times as likely to do so as Democrats, although this result could be explained by the higher prevalence of anti-Clinton misinformation during this period” (Barberá 2). The 2016 election was a precursor to the 2020 election. We saw massive misinformation campaigns in 2016, they were only more prevalent in 2020. Those most susceptible in 2016 may have continued to be susceptible in 2020. But this study was done on a national level. So as we collect more data, we will be able to get a better picture of the groups most susceptible to misinformation in Colorado.

We will extend our project by directly contacting these groups of people. We will conduct semi-structured interviews and rounds of surveys. In order to understand these populations, we must talk to them about their experiences. Our work will extend from these studies as we directly ask these populations about their habits and beliefs stemming from social media. We recognize the boundaries we must set when communicating with possible informants. Their private information will not be shared, but their comments and posts may be used for presentation/analysis/discussion without any identity attached. Participants will be asked a range of questions about their habits on social media and information they are seeing. Along with tweet scraping, our surveys and interviews should give us robust information about Colorado Twitter users.

Methods:

We pursued three methods of data collection. We are focusing on misinformation on Twitter from Colorado Representatives. But in order to study the effects of this misinformation, we needed to find and communicate to the people on the receiving end. That is why we pushed two iterations of our survey, interviewed many Twitter followers of the representatives, and looked at the Tweets at face value. We feel this gave us a robust amount of data about the misinformation sphere within Colorado politics on Twitter. We will discuss how we decided to pursue and conduct our research below.

Surveys

Our initial thought process using surveys as a method of data collection was to gather data from a broad audience in order to narrow our scope for our interviews and second round of surveys. The first round of surveys was created through Google Forms and each member of the group sent out the link through uses of platforms such as Facebook and

Reddit to reach a mass audience. All group members had been a Facebook user, so making a post asking our friends to take a survey was very effective. We also utilized the subreddit r/Samplesize as it was an efficient form of increasing responses in a short amount of time. In our first survey, we used open ended questions with the goal of finding a broader understanding of the people impacted by misinformation. We used the responses from the open ended questions to narrow in the scope for the second round of surveys. We also allowed our respondents an optional part to enter their email address if they wanted to be part of the second round of surveys. The second round of surveys was created through Qualtrics and was fit for our research questions by solely sending it out to Colorado residents. In order to reach politically active and aware Colorado residents, our group was able to join the Facebook groups of both the democratic group and repulican group. The questions asked regarded the truthfulness of representatives' social media feed and the impact it would have on each voter. All respondents' identity was kept anonymous, but in order to ensure they were Colorado residents, we used their IP addresses to verify their geolocation.

Interviews

We decided that talking, face to face with some of these people would give us the best insight into their thinking and political beliefs. But before we reached out to chat with Colorado residents, we conducted needfinding interviews to give us a better idea of our research question. We reached out to fellow colleagues to ask them a couple questions about their social media habits and if they have seen misinformation. We wanted a better understanding from some initial Twitter users. This helped inform us about how we can approach and talk to people in the field. As we gained access to several partisan online political groups, we were able to communicate directly with them. These groups were restricted to Colorado residents. I needed to prove my residency and voter registration in order to be approved to the group. This was extremely important to note because we were challenged when we thought of how we could communicate specifically with a political Colorado population on social media. We were able to distribute our survey in these groups. But the main reason we reached these groups was to talk to these members. These were Facebook groups, but we were looking for Twitter users. We advertised our posts and deliverables to this group by focusing solely on Twitter. This was so we could get Facebook users that also use Twitter. Once we were able to find some members from both sides of the aisle that use Twitter and follow Colorado Representatives, we were able to start data collection.

In order to prepare for these interviews, we did extensive background research. As with our related works section, we needed to understand the populations we are reaching out to. We need to understand how these groups check their information, which demographics are most susceptible to misinformation, and what platforms they use. As

we reached out to many Twitter users on Facebook, we had to adjust our tone and intentions. Our posts in the Colorado Republicans group were met with hostility because a lot of members have left Twitter for certain political reasons. This was something we were anticipating, but it made it more difficult to reach out to these groups even when we had internal access. When we were able to chat over Zoom, email, or phone, we needed to have an interview protocol. This took some time to construct because we wanted to ask similar questions to both political parties. Wording needed to be specific, non-partisan, and not be considered leading. We wanted raw answers about user's routines when confronted with misinformation on a platform. This method of data collection required us to expect the unexpected. This meant answers were startling and responses varied dramatically depending on party and belief affiliation.

Tweet Scraping

To holistically understand the general context of the tweets, two general approaches were undertaken from a qualitative standpoint. The first approach being an understanding of general trends and analysis of both parties verified accounts. The second approach was undertaken to take a look at sentiment and textual understanding of the tweets themselves. In accordance with one another, these comparative approaches allowed us to look at these tweets of interest through multiple perspectives. To specify our tweets of interest within both methods, a specific set timeline was used, as well as the same verified accounts. This allowed each method to utilize and aggregate the same tweets while still pulling differing meta-data and conducting comparative analyses. Our methods for understanding these tweets pulled from multiple sources that particularly specialize in scraping data from Twitter. The libraries used in this research both utilized python specific libraries as well as HTML and XML scraping libraries. The first approach utilized one library called Twint which allowed for easily specified fields, as well as a host of interesting related dimensions. This method aggregated the initial data from the verified Twitter accounts utilizing the timeline of two weeks before and after the 2020 election, with one particular consideration that the verified accounts for Senator Hickenlooper (D) and Representative Boebert (R) were created during this time period and therefore were missing essential information. For this reason these two accounts were dropped from the analysis to better understand those verified accounts that were already developed and active during the specified time period. Beyond this initial omittance, the other active accounts were filtered for the most pertinent features / dimensions, any missing rows or values, and by their partisan value for easier comparison. Then once these tweets and corresponding metadata was organized, a random sample of the tweets was taken from each verified account and aggregated in order to get an unbiased view of the populations. Further analysis and creation of visualizations of this data is discussed in greater detail in the following findings section.

The second method that was undertaken utilized Beautiful Soup, a python specific library that allows easier scraping of HTML and XML files.

Research Findings and Deliverables:

If you conducted any research, please describe the findings. If you have any figures or tables, please include them too. If you conducted interviews, please describe the themes from the interviews and include any quotes to illustrate those themes.

Survey results

The findings from our surveys were very interesting. When we initially created the surveys, we did so with a hypothesis about what kind of responses we were going to see. The expectation was that we would see that in general people have an issue with false or misleading claims about politics being shared on twitter, but that they would have different opinions about what could be considered misinformation. Our findings corroborated this initial prediction. The main theme from our survey was that both liberal and conservative twitter users want to see less misinformation on the platform, but almost all believe that the user is responsible for fact checking for themselves. The majority of respondents also said that their political ideology had either remained unchanged, or they had moved further left since the events of the election. When we consider why this question matters, we can look at the responses to the question “How does the validity/truthfulness of the information posted by your representatives influence your desire to vote for them?” Every respondent who answered this question indicated that they would be less likely to vote for a representative who was shown to spread misinformation. The mistake that was made in survey design that we only realized after the fact was that we didn’t immediately follow this question up with a question asking the respondent to list what kinds of misinformation they have seen spread on twitter. We can infer from other responses like this one ”I think the republicans spewed a massive amount of garbage because they really wanted their fascist in office” that many liberals would have listed the claim that Trump won the 2020 election as misinformation. We missed out on an opportunity to see if the conservative respondents would have listed the same event as misinformation, or if there would have been some responses saying that Biden winning was the lie. From talking with other students, we have also learned that the hard core Qanon conspiracy believers don’t use mainstream media, and in order to survey them, you have to post on the sites that they use, such as parlor and others. In future iterations of this project, we would do exactly that, and reach out to the fringe groups to the best of our abilities. The survey responses help us answer part of the question about how prevalent misinformation is, and how much of an impact it has on the lives of voters. The final conclusion drawn from the surveys was that everyone has an issue with seeing content that they believe is false, but not everyone agrees that it is twitter's responsibility

to moderate the information. This conclusion helped inform our final design recommendation.

Interview Themes

After talking with several Twitter users that are registered Colorado residents and voters, we were able to thematically categorize our notes into a couple specific themes illustrated by their responses.

Censorship

The first theme that was apparent was the topic of censorship. Social media platforms have a lot of power to police their platforms. Hypothetically the platforms can suspend and take down posts that they do not want up. This is something a user gives up as they sign up for the platform. But in recent events, many people feel that what these platforms are doing is considered censorship. From the majority of republican respondents we heard from in the survey and interviews, we got a strong sense that they felt right leaning sentiments were suppressed more than the other side. That is a possibility why some of these users have migrated to other social media platforms. They feel that the platform is against them. This means they will think more things on the platform are misinformation if everything they see is taken down or policed. This was especially evident in responses to the suspension of Donald Trump's Twitter account. Those on the right did not think he should have been banned while those on the left believed he deserved it. This dichotomy between parties is what also heavily influences perception of information on social media. Something we must mention is both sides did talk about censorship in general and that neither group supports it in any form. Our goal in this project is to see how misinformation has pierced certain demographics. Censorship is not a staple of this countries' freedoms, but in some way we want to make sure platforms are policed so the information available is verifiable and reliable. This is a slippery slope we will continue to acknowledge and try to address. As said by one participant, "There should be a fine line between censorship and free speech".

Election Fraud

The most common form of misinformation that persuaded us to pursue the 2020 election was the claim of election fraud. It is common knowledge that election fraud is a harshly punishable crime. There are sporadic cases of election fraud every presidential and congressional election. Out of millions of ballots cast, sometimes a couple may be considered fraudulent. This is just human error, but sometimes people have intentions behind these individual cases. But the reason we chose election fraud as a general theme was because of the prevalence of the idea in the 2020 election. It was known before the election that certain individuals predicted mass-scale fraud because of the increase in mail in ballots. Going into this project, we were expecting to hear about election fraud in

responses. We did not know what they were going to say about it, but knew it would come up because of the magnitude of the topic. We methodically researched how to ask and conduct communication on this topic, as it creates a lot of political division to this day. It was clear that several participants believe that the 2020 election was rigged with fraud. This claim can be animated as one participant mentioned, “It seems more people voted than we have people registered to vote”. They are referring to the state of Colorado which when we do a quick Google search, we see that voter turnout was higher than usual, but not over the registered amount of Colorado voters. There were 3,295,666 ballots counted in Colorado out of a possible 3,793,790 registered voters (CO SoS). Where did this idea about election fraud come from? That is a question we want to try to answer by looking directly at the Tweets. We were able to get a good idea of what accounts these users follow on social media. The same respondent mentioned, “Newt has been saying this. He also says we should demand Colorado legislators to demand a recount”. It is extremely telling that he mentioned the exact source he got this idea from. This user is specifically referring to former Speaker of the House, Newt Gingrich. He did not tell us exactly where he heard this idea from Newt, but it is apparent he got the idea from him. This may be the case for many of his fellow party compatriots. But it is apparent that election fraud was a prominent issue that created misinformation. We want to see if any Colorado representatives pushed a similar idea about a fraudulent election on their social media accounts.

Anger/Partisanship

The last theme that we want to discuss from the interviews is anger. This is another broad theme, but we want to take a look at what drives partisanship, creating division, resulting in anger towards the other side. In simplest terms, with any election, one side will win and the other will lose. One side is happy about the results and one side is obviously not. These events contribute to partisanship. Recent events in politics have created an environment where there is anger towards the candidates, their policies, and the people that actively support them are coming from both sides. This is both in real life and on social media platforms like Twitter. One response that highlights this is when one user said, “We have to stand up to the liberals and big tech companies. They are the true enemies of the people”. Clearly this respondent believes that the left and tech are the ultimate enemies to the right. The mainstream right and left are diametrically opposed on most issues it seems. This will cause an inherent divide, and that has existed for a good portion of this country's history. But partisanship has increased in recent years. The political climate nowadays fosters anger towards the other side. It fuels debates on issues and how issues are framed. This creates misinformation. Partisanship drives misinformation because it creates a notion that the other side is wrong. We saw anger on behalf of users that certain posts and accounts have been flagged for misinformation. They were confused and thus, frustrated as to why the information they saw may have

been false. This turned their anger towards the other side of the political spectrum and towards the platforms. We can highlight this partisanship and anger towards the other side as one respondent said, “the only thing that matters to conservatives is owning the libs. In every way possible. Lies, hypocrisy, projection by any means”. It is clear that the two sides do not like each other. This creates an environment where neither side trusts each other and there is a bigger domain for misinformation to be spread.

How do we want to expand on these themes?
How will they guide our findings and possible proposals?

Tweet Findings / Deliverables

For the coalescence of tweets related to verified Colorado representative’s Twitter accounts, one specific library, Twint, “an advanced Twitter scraping tool written in python”, was utilized. The repository for which can be found here, at their public Github repository. This specific library was used for scraping the tweets we wanted because it allowed for the specification of the start date (October 10, 2020) and the end date (November 11, 2020), exactly two weeks before and after the 2020 election date of November 3, 2020. Another particularly helpful feature within this library was its compatibility with Pandas, which allowed queries to be put directly into dataframes. One such query can be found below for Representative Joe Neguse utilizing Twint and its field specification parameters.

Importation

```
Joe Neguse (D)
https://twitter.com/repjoeneguse?lang=en
2nd district
```

```
b = twint.Config()

b.Username = "RepJoeNeguse"
b.Since = '2020-10-20'
b.Until = '2020-11-17'
b.Pandas = True
```

```
twint.run.Search(b)
```

```
1326968454025515014 2020-11-12 12:21:47 -0600 <RepJoeNeguse> This year's wildfire season 🔥🔥🔥 has been brutal.  And as th
e season moves later into the year, it is critical our contracts with hotshot fire-fighting crews are extended as well.  Ou
r letter with @RepKatiePorter to DOI &amp;amp; USDA is an important first step to making that happen!  https://t.co/malh5IULwP
1326942878212481026 2020-11-12 10:40:10 -0600 <RepJoeNeguse> Last night we lit the #Boulder Star in remembrance of all the
veterans who have sacrificed so much for our country. The star has served as a beacon of hope and our resiliency since it w
as first displayed 73 years ago, and it was an honor to join the ceremony this year.  https://t.co/5A692YyUEX
1326637253318209538 2020-11-11 14:25:43 -0600 <RepJoeNeguse> The opioid epidemic has taken far too many lives. And yet the
current deal between the Department of Justice and Purdue Pharma would not result in justice for those who have been hurt.
Proud to join @RepKClark's letter below imploring the DOJ to hold them fully accountable.  https://t.co/M4cl5ZOpni
1326531255899316225 2020-11-11 07:24:31 -0600 <RepJoeNeguse> Today we honor those who have so bravely fought for and served
```

After the queries were called for each verified account created during the time period, the following dataframes were filtered and cleaned for unneeded dimensions. Primarily, data regarding the tweet itself, and any meta-data such as amount of likes, retweets, or replies were used in order to understand the popularity and influence surrounding the accounts. Once each of the usable representatives was filtered and cleaned for useful information, they were aggregated by multiple iterations of samples taken from each account's tweet

related data. Averages were then taken from the multiple samples of each account and averages to give a more unbiased view of their activity. Below you will find dataframes for both the unfiltered / filtered dataframes, and dataframes, by party, of the sampled averages for each representative.

Unfiltered / Filtered Dataframes

	id	conversation_id	created_at	date	timezone	place	tweet	language	hashtags	cashta
0	1327694252978139138	1327694252978139138	1.605382e+12	2020-11-14 12:25:51	-0600		Happy Diwali to all those who celebrate. Wishi...	en	[]	[]
1	1327362573696643074	1327362573696643074	1.605303e+12	2020-11-13 14:27:53	-0600		Colorado's public lands are more than just bea...	en	[pawact]	[]
2	1327339845815996417	1327339845815996417	1.605297e+12	2020-11-13 12:57:34	-0600		Abortion care is health care, period. No cli...	en	[gagrule]	[]
3	1326987342532501505	1326987342532501505	1.605213e+12	2020-11-12 13:36:51	-0600		So excited that there will be a little piece o...	en	[]	[]
4	1326903231256080388	1326903231256080388	1.605193e+12	2020-11-12 08:02:37	-0600		THIS IS IMPORTANT - the CDC now says that wear...	en	[wearamask]	[]

	date	tweet	username	name	retweet	nlikes	nreplies	nretweets
0	2020-11-14 12:25:51	Happy Diwali to all those who celebrate. Wishi...	RepDianaDeGette	Rep. Diana DeGette	False	45	3	5
1	2020-11-13 14:27:53	Colorado's public lands are more than just bea...	RepDianaDeGette	Rep. Diana DeGette	False	34	1	2
2	2020-11-13 12:57:34	Abortion care is health care, period. No cli...	RepDianaDeGette	Rep. Diana DeGette	False	49	9	12
3	2020-11-12 13:36:51	So excited that there will be a little piece o...	RepDianaDeGette	Rep. Diana DeGette	False	64	5	4
4	2020-11-12 08:02:37	THIS IS IMPORTANT - the CDC now says that wear...	RepDianaDeGette	Rep. Diana DeGette	False	38	5	8

Sampled Averages by Party

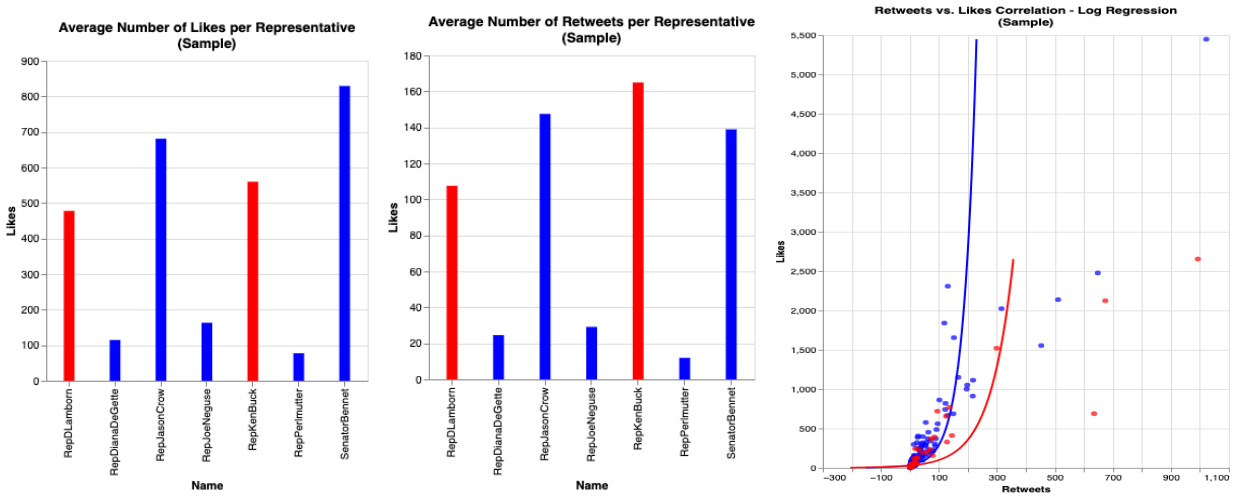
index	nlikes	nreplies	nretweets	name
RepKenBuck	559.909091	104.681818	164.909091	RepKenBuck
RepDLamborn	477.812500	97.312500	107.500000	RepDLamborn

index	nlikes	nreplies	nretweets	name
RepDianaDeGette	114.620690	15.448276	24.551724	RepDianaDeGette
RepJoeNeguse	163.176471	8.441176	29.117647	RepJoeNeguse
RepJasonCrow	680.750000	104.325000	147.450000	RepJasonCrow
RepPerlmutter	77.379310	11.068966	11.896552	RepPerlmutter
SenatorBennet	829.123077	61.492308	138.846154	SenatorBennet

Some particularly interesting findings after conducting this analysis was that not only were both parties represented by one or more popularized accounts, but that there was

also a great amount of influence those popularized accounts circulated. As shown in the visualizations below there were great amounts of likes and retweets within two accounts for the Republican party and two for the Democratic before and after the election. This generally reiterated connotations the state of Colorado represented through its overwhelmingly bi-partisan population centers. Furthermore, for both parties accounts there was a positive correlation between the amount of retweets each account procured and the amount of likes. This generally means that both types of partisan accounts found themselves receiving a proportional amount of likes to the amount of retweets they received. For specific tweets within popularized accounts, this looks like a significant amount of influence or circulation throughout online partisan social circles.

Visualizations



Twint Repository
https://github.com/twintproject/twint

How do these findings connect?
What does the culmination of all this data reveal about our research question?

Twitter does not have a robust method of labelling and rooting out misinformation on their platform. In a quote on the Twitter Blog, a representative of the company wrote “We are not the arbiters of truth on our platform.” We are paraphrasing a longer post here but in essence they said that there are too many posts to have staff verify, and they rely on experts in the community to call out false claims in the comments under these posts, and to report them for twitter to review as best they can.

As a result of our surveys, we see that people encounter misinformation on platforms like Twitter all the time. Respondents also told us how they try to deal with it, by

independently fact checking it, ignoring it, or even asking a family member to help them identify if the information is valid. These methods of relying on yourself, or a trusted family member helped inform our final design recommendation.

Looking directly at the posts from Colorado's representatives, we did not find concrete examples of blatant misinformation about the 2020 presidential election. We found some possible semi-misleading tweets, but nothing that was flagged by Twitter as misinformation. This does make sense with how twitter treats misinformation on their platform. However, since we didn't find any tweets containing misinformation, we were able to answer our research question. Colorado representatives did not spread misinformation from our specific timeline around the 2020 election. We do have to say, our method for collecting tweets doesn't collect deleted tweets, so if there were misleading or false claims made that a representative later deleted, we were not able to see those. However, if this deleting of misinformation did occur, it does indicate that Colorado's representatives do not want to spread misinformation. If they were ok with or actively trying to spread misinformation, there is a strong possibility that tweets containing false claims would have been recovered. We cannot prove this with any degree of certainty, because the absence of something is not evidence to prove it ever existed.

Interviews showed us an insight into the populations. With more personal responses about beliefs and habits, we can see how misinformation has infiltrated these groups. Ideas about election fraud, censorship, and blatant partisanship has proved to us that some groups were susceptible to misinformation in general. This is apparent going into the project, but we have a better idea of who it is that is believing this false information.

Representatives are elected by the people to represent the people. In the past, elected officials such as members of congress both local and national, as well as governors and mayors, leaders all the way to the presidents themselves have been judged based on their words and actions on the public stage. Since the early 2000's, that public stage has extended into the virtual world, and we need to hold elected officials accountable for their conduct in the real world, and in the virtual one. Our surveys indicate that many people already consider the truthfulness of their representatives online when deciding to vote for them, so as a recommendation to those representatives, we suggest treating tweets and social media posts the same way they treat a press conference or town hall.

Our final design recommendation that we came to as a result of all our findings is that Twitter should implement a community peer review feature. Our idea is that it could function similar to the reddit Upvote/Downvote feature, but instead be based on truthfulness. After a certain number of people report a tweet for misleading information,

Twitter could add a banner on the bottom automatically where users then upvote or downvote the tweet. This banner stays as an indication of how the community feels about the information, but the tweet doesn't get removed. This is because both twitter, and a lot of its users indicate that they do not want to be deciding what is and isn't shown, even if it is considered misleading. In future iterations of this project, we could create a prototype of the system, and have users give us feedback on how they feel about it, and how well it works.

Conclusion:

Next Steps

If we had more time to complete this project we could have expanded it. We were able to obtain a lot of information, but given a longer timeline for collecting Tweets and talking to users would have given us more robust findings. Over the short timeline, we were not able to get all the information we would have liked. Extending this timeline would definitely give us a stronger idea if Colorado representatives actually spread misinformation. We could take this a step further and look at even more representative Tweets. Our initial research question before we narrowed it was going to take a look at all 100 sitting U.S. Senators. For the timeline of this project, choosing just Colorado representatives saved us time and gave us a specific population to look at. Another way we could expand on this preliminary research project is taking a look at other social media platforms. We could figure out what other platforms these representatives use on a daily basis and scrape those posts. This would allow us to further expand into representatives personal profiles because there was a stark contrast in popularity or circulation surrounding their varying accounts when conducting our research. We would hope for more diversity of posts/information from other platforms, but this is just an idea. An additional way in which we could expand our research would be to compare verified representatives amongst varying states. By observing differing popularized representative accounts in largely partisan states, this would allow us to compare our generally bi-partisan findings and offer more insightful socio-political conclusions. This further expansion of our research would primarily revolve around a specific social media platform. Some people only use certain social media applications, and this was a large factor throughout our research. There are many next steps we could pursue from this project.

Reflect on team and project management

We all met in a brainstorming breakout room and discussed what topics we found interesting. Us five really liked the idea of trying to find and analyze misinformation. We decided that Twitter would be the best application to use because it is relatively easy to scrape posts. We needed to make sure we had multiple methods of data collection so we

could synthesize all our information together. We decided to conduct rounds of surveys, interviews, and Tweet scraping. We also divided ourselves amongst these strategies. Dylan and Shota helped formulate survey questions and analysis of those surveys. Alex reached out and contacted directly with Colorado residents. Finally, Hayden and Connor scraped posts from Twitter so we could have the raw data. We went separate ways when collecting the data, but came together to make sure we had sound protocols for asking questions, communicating with audiences, and reporting the data. It was effective for us to divide the data collection like this because it catered to our skills. Dylan and Shota have conducted many surveys with different technologies and topics. Alex has completed many extensive research projects that involved contacting specific communities. And Hayden and Connor are more technically gifted and were able to effectively use many Python libraries to obtain posts. This combination helped us split the work up evenly and keep everyone busy. Overall, we were able to complete this project with a couple minor hiccups.

What advice would you give to future INFO seniors taking this class?
This was a very interesting class. We were all excited to take a class where we could focus on one project for the whole semester. We knew we could dedicate a lot of time and resources to this and we could delve deep into the topic at hand. Because you are allowed the freedom to choose your group and research question, do something you want to. You will be more invested in the project and hopefully put more work into it.

Do not procrastinate. If you are taking this class you are a senior. You may be graduating after this class or soon after. You will have a lot of other work in other classes. This is pretty standard in college, you are four years into it now. But make sure you pace yourself throughout the semester. Create milestones for your project. That is what we did. We were able to effectively communicate when we wanted certain aspects and deliverables of the project done. This will be one of your biggest projects of your time at CU. Do not wait to the last minute to craft something together. Gather your data, ideas, and group mates a lot so you can continue to document your project progress. Hopefully you will be proud of your work if you do something along these lines.

Works Cited

Barberá, P.. “Explaining the Spread of Misinformation on Social Media: Evidence from the 2016 U.S. Presidential Election.” (2018).

Election Night Reporting, 4 Mar. 2021, results.enr.clarityelections.com/CO/105975/web.275533/#/summary.

Kouzy R, Abi Jaoude J, Kraitem A, et al. (March 13, 2020) Coronavirus Goes Viral: Quantifying the COVID-19 Misinformation Epidemic on Twitter. Cureus 12(3): e7255. DOI 10.7759/cureus.7255

O'Brien, Matt. “Did Social Media Actually Counter Election Misinformation?” *AP NEWS*, Associated Press, 5 Nov. 2020, apnews.com/article/social-media-election-misinformation-632a5d93a6cc3ff37311a641d6bf5a1.

“2021 Storming of the United States Capitol.” *Wikipedia*, Wikimedia Foundation, 29 Apr. 2021, en.wikipedia.org/wiki/2021_storming_of_the_United_States_Capitol.