STEM Fellowship Undergraduate Big Data Challenge 2021

Are You Spreading Misinformation? Analyzing How Influential Twitter Users Contribute to the Spread of COVID-19 Information

Bizhan Alatif[1], Jennifer Tram Su[1], Maggie Wang[2], and Sarina Xi[2]

[1]McGill University
[2]University of Toronto

July 4th, 2021

Abstract

When COVID-19 was declared a pandemic in early 2020, a virtual pandemic of misinformation had gone undeclared. While some of these misinformed claims may be tame, others could lead to the loss of life (e.g. Hydroxychloroquine curing COVID-19) and behavioral changes with larger ramifications (e.g. Bulk buying hand sanitizer). Therefore, detecting and removing instances of misinformation could reduce the consequences that arise from a misled public. In particular, misinformation tends to pervade social media platforms. Its convenience allows users to easily consume and spread information to others through posting. As a result of this accessibility, many have turned to these outlets as their primary source of news [1]. In 2020, the three most used social media platforms were Facebook, YouTube, and Twitter. While Facebook dominates as the go-to platform for general use, over half of Twitter's user base (59%) is reliant on the platform for news [2]. Furthermore, Twitter is known as the platform that houses the personal thoughts of many well-known figures. As opposed to Facebook or YouTube, where these interactions are scarce, users can directly engage with these influencers and easily pass on the content produced by these figures for the user's own network to see. Thus, Twitter's social structure allows information to easily propagate between influencer and user, and amongst users themselves. Accordingly, this paper examines the correlation between influential Twitter accounts – right and left-leaning news sources, health organizations, celebrities, and right and left-wing politicians – and the spread of misinformed claims on Twitter. Being higher-level nodes in their social networks, these Twitter accounts are selected as their large follower counts allow for information to easily disseminate to a wider audience [3]. Visualizing these interactions can provide insight into the patterns and trends allowing for the identification and cessation of misinformation spread. This paper concentrates on misinformation, predominantly in the USA, during the early stages of COVID-19 by comparing tweets from well-known Twitter accounts to datasets consisting of "fake" and "true" claim tweets. Additionally, engagement metrics are used to gauge how much traction tweets from different influential Twitter accounts garner. Further work with sentiment analysis is also conducted to explore how users respond to tweets from influential figures - whether they amplify or dissuade information based on sentiment (e.g. feeding into fear-mongering) and whether sentiment acts as a motivator in the spread of information. With this, timelines of tweet frequency, engagement, and sentiment are used to determine the correlations, via application of the Kendall Rank Correlation Coefficient and Mann-Kendall significance test, between influential Twitter groups and general fake and true tweets. The results show that the most statistically significant relationship found is between the distribution of real tweets and average likes and retweets from health organizations. Other social network nodes do not have concrete relationships with fake and real tweets. It is also observed that sentiment is not a driving factor, however, additional exploration into the influence of sentiment is recommended.

Keywords
COVID-19, Pandemic, Fake News, Misinformation, Tweets, Social Media, Twitter

1 Introduction

Since the introduction of social media, people have been able to connect with one another in an instant. While this leads to information quickly spreading amongst individuals and groups, it also leads to the swift spread of misinformation. Furthermore, social media has overtaken mainstream news networks as the predominant news source, with 86% of Americans using social media over other news outlets [1]. Since social media content is heavily dependent on user engagement (e.g. retweets, sharing, and posting), false claims can be easily proliferated by extensive engagement, leading to misinformed claims reaching a wider audience. Additionally, as a result of the current over-reliance on social media for news, people could easily be misled when encountering misinformation. Recently, under the influence of the COVID-19 pandemic, dozens of false claims and headlines surfaced, with some of the notable claims connecting COVID-19 to 5G/radio-waves [4], proposing cleaning agents as a preventative measure, and that the use of hydroxychloroquine can cure COVID-19 [5]. This surge of misinformation impacts the livelihoods and health of those led to believe them, as well as the general public that lives alongside them. In response to this, the World Health Organization resurfaced a previously coined term, "infodemic," to describe the overabundance of misinformation emerging online and offline in their efforts to bring awareness to this misinformation pandemic [6].

In 2020, the three most used social media platforms were Facebook, YouTube, and Twitter. While Facebook dominates as the go-to platform for general use, over half of Twitter's user base (59%) is reliant on the platform for news [2]. Furthermore, Twitter is known as the platform that houses the personal thoughts of many well-known figures. As opposed to Facebook or YouTube, where these interactions are scarce, users can directly engage with these microblogging influencers and easily pass on the content produced by these figures for the users' network to see.

Similar to other social media platforms, Twitter has a constantly evolving algorithm that attempts to show the most relevant tweets to users. This is accomplished by ranking tweets based on various criteria (engagement, user activity, tweet age, and special media usage - GIFs, emojis, etc...) [7]. Most notably, Twitter's algorithm favors tweets that garner significant engagement, which is tallied in the form of likes, replies, retweets, and quote retweets (note that quote retweets are similar to retweets, with the added feature that users can retweet an original tweet and post a comment attached to the respective retweet [8]). Likes and replies are similar to all other social media: numerical values describing how much direct engagement a tweet has. While modeling the propagation of tweets based on the metric of likes and replies is difficult as they are characteristics of an algorithm, retweets and quote retweets directly propagate information to a new network of individuals; however, retweets show a sudden cascade of informational propagation which exponentially dies out quickly [9]. Early on, once a tweet is retweeted, it is promptly retweeted by other users, indicating rapid informational diffusion [10]. In fact, users' retweeting shows preferential attachment as tweets with a higher number of retweets will likely be retweeted more often [11]. Therefore, it's desirable to analyze the engagement of tweets as they are directly related to the source tweet's (mis)informational propagation. In this analysis, a focus is given to retweets and likes since likes are representative of the general traction for a tweet and retweets aid with (mis)information propagation.

This analysis considers various large sub-groups of "influential" Twitter accounts - health organizations, left/right-wing news sources and political figures, and celebrities. Since the largest informational cascades tend to center around influential users (high follower accounts or previously influential individuals) [12], the goal is to determine which sub-groups most greatly influence the spikes in (mis)information in the form of tweet frequency. However, since this cascade effect is unreliable on an individual basis and only reliable with large numbers of influencers [12], it is also crucial to pool together numerous individual Twitter accounts into each sub-group. Similar to social network analysis (SNA) [13], this paper views influential figures' accounts as social network nodes. The social structure of Twitter allows for users to connect and engage over common ideologies. With larger network nodes amassing large follower counts, the information they release easily disseminates along its many edges. Therefore, according to Dynamic Social Impact Theory (DSIT), higher-level nodes in central positions of a network are perceived as more influential and powerful [14]. In the context of COVID-19, these high nodes can contribute to the spread of true and fake news. To better understand how misinformation spreads on social media, the correlations between influential social network nodes and the spread of fake and real tweets are studied. Using data sets consisting of tweets between March and May 2020, correlations were calculated between the spread of misinformation and social network nodes based on their tweet frequency, average number of likes, and average number of retweets.

Another aspect is sentiment analysis, which is a method focused on identifying and quantifying the emotions in language [15]. Based on the findings of Jimenez-Zafra et al. [15], negative sentiment tends to correlate with increased retweeting frequency. If this holds, misinformation should have a higher probability to spread if the source tweets contain negative sentimental language. Furthermore, Nanath and Joy [16] found that tweets are more likely to be retweeted if they come from influential entities, contain negative emotions, or have optimistic information. As a result, it is significant to further study how the sentiment of tweets may impact the virality of information shared over Twitter. This can then be used to answer the question "What sentimental language is typically associated with misinformation".

To counter this spread of misinformation, several solutions are being proposed and in the works. For instance, Lanius et al. [17] found that flagged tweets reduce the participant's attitude towards it, with some participants changing their opinion on the matter. Furthermore, according to Blankenship [18], Twitter must introduce readily available fact-checking systems for users, providing them a means of gauging the credibility of what they read. Additionally, identifying bot accounts is necessary as they intentionally spread fake news. Twitter must, therefore, develop mechanisms and tools to automate this identification [18].

Our analysis takes a different route compared to the aforementioned ones. We seek to determine answers to the question of "Which influencers are more likely to spread misinformation or information" and move towards calling for a solution based on these results.

2 Materials & Methods

Table 1 provides examples of the Twitter accounts considered as influential social network nodes in different groups. The three main groups considered were "Health Organizations", "News Outlets", and "Individuals". The latter two groups were further subdivided based on their political views into "Left" and "Right". Accounts for the "Health Organizations" category were restricted to governmental and educational establishments related to the medical field. The "Left" and "Right" news outlets were determined based on a study classifying the political-leaning of each news network [19]. "Political Left" and "Political Right" individuals were chosen strictly based on their political affiliations (e.g. Republican/Democrat). Finally, sixty-three "Celebrities" were taken out of the top hundred most followed Twitter accounts, excluding politicians and groups (e.g. Barack Obama, Manchester United, One Direction). Tweets made after January 1, 2020, from accounts in each group were scrapped using Twint [20], where relevant tweets were extracted by filtering for case-insensitive COVID-specific substrings in 1.

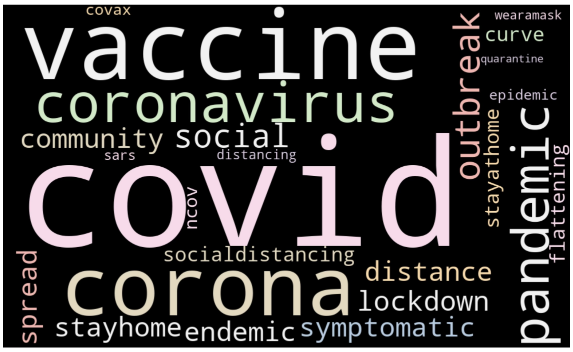

Figure 1: COVID Substrings

Data containing tweets that contribute to the spread of COVID-19 misinformation are collected from various open-source datasets. The specific sources of datasets are shown in Table 2. These datasets are already labelled as "real" or "fake" news. The two categories were separated for independent comparison with each Twitter group. To ensure consistency of the distribution of all datasets, only tweets from March to May of 2020 are used.

A timeline of tweet frequency as well as two types of engagement timelines were created to assess the spread of misinformation. The two categories of engagement include the average number of likes and retweets per tweet over time (i.e. $\frac{Likes}{Tweets}$ and $\frac{Retweets}{Tweets}$). While replies are another significant form of platform engagement, this metric is neglected as not all data sets provided reply counts.

Further investigation was done through sentiment analysis. The tweets were processed to remove hyperlinks, "RT"s, punctuation, special characters, and converted to lowercase. Using the Python library, textblob[21], the processed tweets for each group were scored on a polarity scale from -1.0 to 1.0, representing negative and positive respectively, as well as 0.0 for neutral.

All timelines were created using the Python libraries Pandas [22], Matplotlib [23], and NumPy [24].

Examples of Influential Twitter Accounts	
	Examples
Health Organizations	CDCgov, WHO, HarvardChanSPH, JohnsHopkinsSPH
Right-Leaning News Outlets	FoxNews, Real Daily Wire, The Blaze
Left-Leaning News Outlets	ABC News, CNN, Guardian, VICE, TheDailyShow
Political Right	Boris Johnson, Dave Rubin, Donald Trump
Political Left	Andrew Yang, Hillary Clinton, Justin Trudeau
Celebrities	Adele, Bill Gates, Elon Musk, Jimmy Kimmel

Table 1: Examples of Influential Twitter Accounts

Finally, tests to measure the correlation between the frequency, sentiment, and engagement timelines are performed using the Kendall Rank Correlation Coefficient (KRCC). This coefficient is a non-parametric hypothesis test that provides insight into the strength of dependence between two variables [25] - in this case, the tweet frequencies and engagement metrics between two groups, and similarly, the sentiment frequencies and engagement metrics between two groups. To test for their significance, their p-values are calculated based on the Mann-Kendall significance test, with the null-hypothesis being a KRCC of 0 (no dependence). After binning the data in intervals of three-days, the KRCC value, τ_b, is calculated by plotting the binned y-axis values of the timelines of each influential Twitter groups against the datasets of real and fake tweets and performing the following operation with the set of (x,y) pairs: [25]

$$\tau_b = \frac{n_c - n_d}{(n * (n-1)/2)}$$

Where n_c is the number of concordant pairs where for $i < j$, $x_i < x_j$ or $y_i < y_j$, n_d is the number of discordant pairs if the pair is not concordant, and n is the total number of pairs.

The p-values are calculated using the kendall-tau function in Python's Scipy [26] library with the following theory behind it [27]:

$$z_b = \frac{3(n_c - n_d)}{\sqrt{n(n-1)(2n+5)/2}}$$

Open-source Data Sets Used for Analysis
TathyaCov: Detecting Fake Tweets in the times of COVID 19 [28]
Cross-SEAN: A Cross-Stitch Semi-Supervised Neural Attention Model for COVID-19 Fake News Detection [29]
COVID-19-FAKES: A Twitter Dataset for Detecting Misleading Information on COVID-19 [30]
Dataset for COVID-19 Misinformation on Twitter [31]
CMU-MisCov19: A Novel Twitter Dataset for Characterizing COVID-19 Misinformation [32]

Table 2: Open-source Data Sets Used for Analysis

where for a two-tailed hypothesis test, the p-value is twice the z_b statistic.

3 Results

Figure 2 shows the various frequency timelines from March 1^{st} to May 25^{th} 2020 separated into each category of influential Twitter accounts. The two last plots represent the spread of fake and real tweets collected from open-source datasets. The blue line represents the daily tweet frequency fluctuations, while the histogram is binned every three days.

Figure 3 and Figure 4 represent the engagement timeline plots, while Figure 5 represents the tweet sentiment timeline plot - all over the same period as Figure 2 (tweet frequency). Figure 3 shows the average daily likes timelines and Figure 4 shows the average retweets timelines. Figure 5 shows the frequency of negative, neutral, and positive tweets per day. Similar to Figure 2 (tweet frequency), a histogram binned every three days is used for Figures 3 and 4 (average likes and retweets). Note that since the quantity of data across all categories varies and no normalization was applied, the magnitudes of the vertical axes differ.

From these timelines, the Kendall Rank Correlation Coefficient (KRCC) and the Mann-Kendall significance test p-value are determined to evaluate the correlation between pairs of influential social network nodes and false and true tweets. Tables 4 and 5 show the calculated pairs of KRCC and p-values for tweet frequency and the two engagement metrics (average likes and average retweets). Tables 6 through 7 show the calculated pairs of KRCC and p-values for the sentiment frequencies between influential nodes and false and true tweets. The KRCC values between the sentiment frequencies and the two engagement metrics can be found in the appendix.

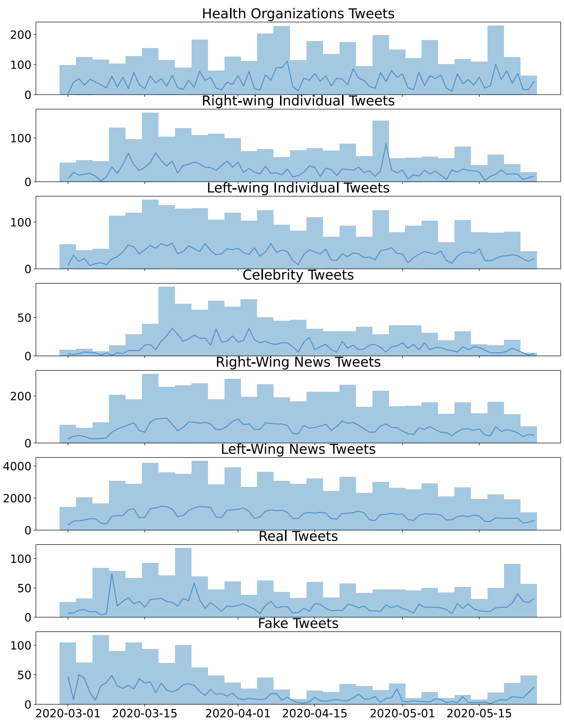

Figure 2: Timelines for Tweet Frequency

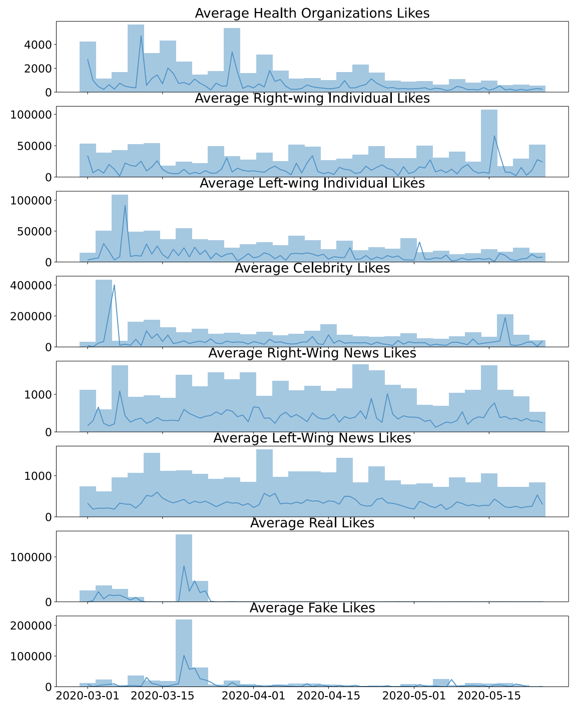

Figure 3: Timelines for Average Likes

Table 3: Tweet Frequency Timeline Correlation KRCC and Mann-Kendall P-Values

		KRCC	P-Value
Health Org.	Real	0.1053	0.5061
	Fake	-0.0798	0.5476
Political Right	Real	0.2354	0.0745
	Fake	0.2129	0.1065
Political Left	Real	0.3192	0.0161
	Fake	0.2217	0.0945
Celebrities	Real	0.1197	0.3670
	Fake	0.0150	0.9102
News Right	Real	0.1687	0.2016
	Fake	0.1214	0.3576
News Left	Real	0.2602	0.0487
	Fake	0.1980	0.1332

Table 4: Average Likes Timeline Correlation KRCC and Mann-Kendall P-Values

		KRCC	P-Value
Health Org.	Real	0.3153	0.0162
	Fake	-0.0690	0.6156
Political Right	Real	0.0246	0.8672
	Fake	-0.1084	0.4233
Political Left	Real	0.3300	0.0117
	Fake	0.2956	0.0246
Celebrities	Real	0.1084	0.4233
	Fake	0.3990	0.0020
News Right	Real	0.2414	0.0685
	Fake	-0.0887	0.5149
News Left	Real	0.0591	0.6689
	Fake	0.0246	0.8672

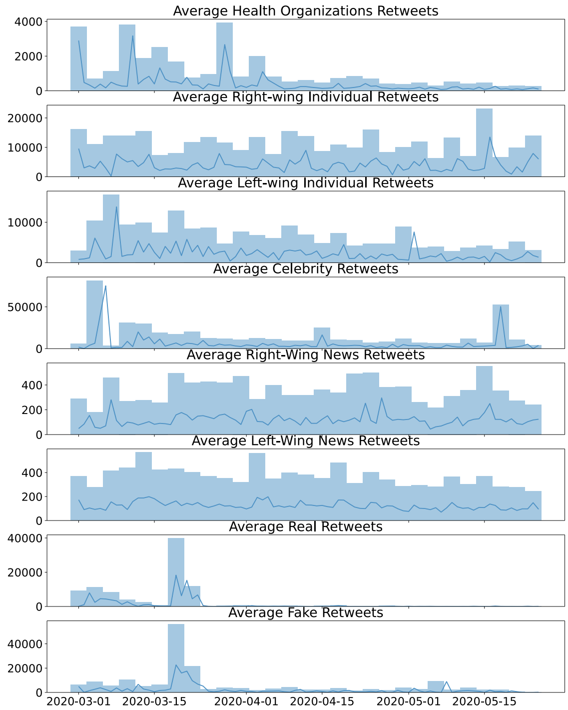

Figure 4: Timelines for Average Retweets (placeholder)

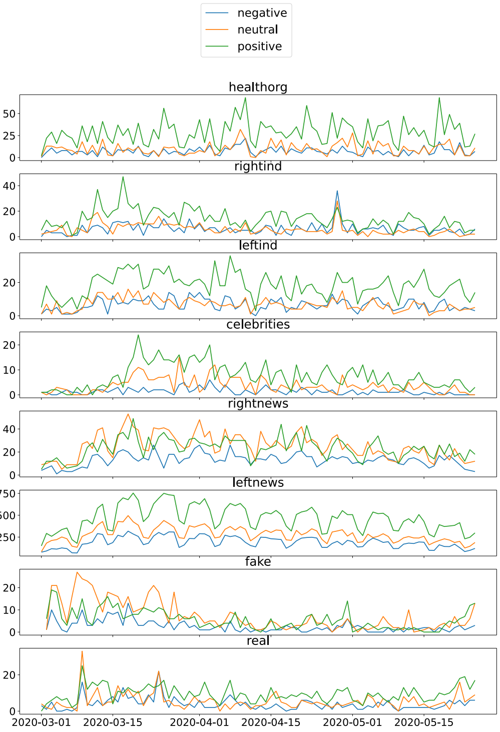

Figure 5: Timelines for Tweet Sentiments

Table 5: Average Retweets Timeline Correlation KRCC and Mann-Kendall P-Values

		KRCC	P-Value
Health Org.	Real	0.4138	0.0013
	Fake	-0.0739	0.5896
Political Right	Real	0.0099	0.9555
	Fake	-0.0197	0.8965
Political Left	Real	0.4729	0.0002
	Fake	0.3842	0.0031
Celebrities	Real	0.3251	0.0131
	Fake	0.1970	0.1394
News Right	Real	0.0394	0.7804
	Fake	-0.0591	0.6689
News Left	Real	0.3153	0.0162
	Fake	0.2562	0.0528

Table 6: Sentiment Real Tweets Timeline Correlation KRCC and Mann-Kendall P-Values

		KRCC	P-Value
Health Org.	Pos	0.1357	0.3096
	Neu	-0.0939	0.4855
	Neg	0.0945	0.4847
Political Right	Pos	0.4086	0.0024
	Neu	0.0153	0.9099
	Neg	0.2843	0.0361
Political Left	Pos	0.3506	0.0089
	Neu	0.2013	0.1362
	Neg	0.2879	0.0330
Celebrities	Pos	0.0761	0.5717
	Neu	-0.0179	0.8949
	Neg	0.1571	0.2602
News Right	Pos	0.3001	0.0251
	Neu	0.1310	0.3275
	Neg	0.2639	0.0499
News Left	Pos	0.3300	0.0131
	Neu	0.1330	0.3186
	Neg	0.3530	0.0080

Table 7: Sentiment Fake Tweets Timeline Correlation KRCC and Mann-Kendall P-Values

		KRCC	P-Value
Health Org.	Pos	-0.1945	0.1427
	Neu	-0.0652	0.6249
	Neg	-0.1535	0.2631
Political Right	Pos	0.2116	0.1139
	Neu	0.2965	0.0264
	Neg	-0.0235	0.8645
Political Left	Pos	0.0676	0.6117
	Neu	0.1887	0.1581
	Neg	-0.0467	0.7331
Celebrities	Pos	-0.0403	0.8069
	Neu	0.1791	0.1762
	Neg	-0.0620	0.6495
News Right	Pos	0.0325	0.8069
	Neu	0.1791	0.1762
	Neg	-0.0620	0.6495
News Left	Pos	0.0496	0.7072
	Neu	0.2255	0.0875
	Neg	-0.0025	0.9849

4 Discussion

After producing the timelines, KRCC values are used to evaluate the correlation as they are a distribution-free test that can measure the strength of dependence between two variables [33]. As the distribution of the timelines produced are unknown, KRCC is a preferable way to assess their correlations. Its limitation of assessing between only two variables also fits into the problem definition of this project as we do not test for correlations between different influential social nodes as they are linked poorly. [34].

The KRCC between tweet frequency and real and fake news is interpreted as the correlation between the user's posts and the spread of either form of information. The KRCC between engagement metrics of likes and retweets, and real and fake tweets is interpreted as a measure of similarity between the two audience groups. It is used as an additional metric for how each group contributes to the spread of misinformation, as different audience demographics interact differently with social media [35]. Similarly, the KRCC between the sentiment frequencies of the influential nodes and real and fake tweets is interpreted as the correlation between the sentiment of the user's posts and spread of either form of information. The sentiments of interest are 'negative', 'neutral', and 'positive', with emphasis on 'negative' and 'positive' during analysis. Furthermore, the KRCC between the sentiment frequencies and engagement metrics is interpreted as a measure of how the audience groups respond to each sentiment. In other words, to explore whether different influential nodes tend towards certain sentiments, and whether their audience further fuels those sentiments with higher engagement. However, it should be noted that these sentiments have no indication towards agreement or disagreement, nor are they an indication of subjectivity; they are only used to evaluate the degree of interaction.

Likes and retweets were measured separately as they measure different forms of engagement. Retweets represent a higher level of engagement compared to likes, as users must consciously decide to share the information [36]. Furthermore, likes are a passive activity, and this form of engagement contributes less to the spread of content.

To go about defining different levels of significance interpreted from the KRCC, this paper follows the proposed scale: [37]

- $|\tau_b| = 0.07$ indicates a weak association
- $|\tau_b| = 0.21$ indicates a medium association
- $|\tau_b| = 0.35$ indicates a strong association

4.1 Tweet Frequency

Overall, health organizations show a consistently moderate correlation across all three metrics with real news, and left-wing political figures show a moderate correlation with real and fake news. Looking at frequency alone, news groups and political figures generally have a moderate correlation with both types of news, while health organizations and celebrities have a low correlation. However, for most groups, these results are not corroborated by the average likes and average retweets.

From the KRCC and p-values, one can see that for health organizations, the tweet frequency KRCC is a negative value of -0.0798 for its correlation to fake tweets, while it has a higher tweet frequency KRCC of 0.1053 for true tweets in comparison as shown in Table 3. The p-value for fake tweet frequency of 0.5476 in Table 4 is high, which indicates that health organizations have a high probability of achieving the null-hypothesis: not correlated to fake tweets. These trends are further reflected in the engagement correlations from Tables 5 and 7, where average likes and retweets both have a much higher correlation to true tweets compared to fake tweets. This could suggest that groups following real news are more likely to also follow health organizations.

From Table 3, the tweet frequency of politically left-leaning news sources has a moderate and moderately strong correlation with real and fake tweets, respectively. Similarly, the relationship between average retweets and real/fake retweet engagement is shown to have moderate and strong correlation for real and fake tweets, respectively. Finally, average likes appear to show no discernible correlation. However, in analyzing the p-values from Tables 3 to 5, the only statistically significant values are the relationship between left-wing news sources and true tweet frequency with a KRCC of 0.1687 and a p-value of 0.0487 as shown in Table 3, and with true average retweet frequency with a KRCC of 0.3153 and p-value of 0.0162 in Table 5 (all other relationships would then require further evidence to support any relationship). Therefore, taking significance into account, real tweets may have a more correlated distribution to left-wing news sources relative to fake tweets. The average likes metric for left-wing news sources does not have any significant trends. The statistically significant correlations between tweet frequency and average retweets with that of real tweets suggests the possibility that left-wing news outlets may impact the spread of real news. Although the correlations with fake tweets are not significant, their strengths are comparable to the correlations with real tweets. Further investigations must be made to draw conclusions between left-wing news and fake tweets, as it is possible that left-wing news outlets effectively spread both real and fake news.

When examining Table 3, tweet frequency from politically left-wing individuals appears to have moderately strong correlations with real tweets of 0.3192, as opposed to a moderate correlation with fake tweets of 0.2217. These former correlations are shown to be statistically significant with a p-value of 0.0161 for true tweets. From Tables 4 and 5, there is a slightly stronger correlation between engagement and real tweets compared to fake tweets. For average likes, the correlation with real tweets is 0.3300 compared to 0.2956 for fake tweets. For average retweets, there is a KRCC of 0.4729 with real tweets and a lower, but still strong correlation of 0.3842 with fake tweets. All of these values are statistically significant with p-values below 0.05. Considering the tweet frequency, the preference for real tweets suggests the possibility that politically left-wing individuals may impact the spread of real tweets more than fake tweets and propagate the same information through shared audience bases. However, the strong correlations between the engagement metrics and fake tweets indicate that there is still considerable overlap between left-wing individuals' tweets and fake tweets. While fake tweets may not spike at similar times as left-wing tweets, they may still propagate the same information from the influencers' tweets.

The correlations for celebrity individuals have varying results across all three metrics of comparison for KRCC values as shown in Tables 3 to 5. Focusing on average retweets, there is a strong correlation with real news of 0.3251, shown to be statistically significant through a p-value of 0.0131. However, when analyzing average likes, fake tweets have a strong, statistically significant relationship, a KRCC of 0.3990, with celebrities, while the weak correlation with real tweets is not significant. Neither of these observations are supported by the tweet frequencies, which is insignificant for real tweets. However, the low KRCC and high p-value between celebrities and fake tweets suggests that they are unlikely to propagate fake news. Altogether, this could imply that celebrities' audience members engage with both real and fake tweets, but celebrities themselves do not propagate or spread either type of tweet. The mismatch in tweet type may be a result of the different types of celebrities included in this analysis, and consequently, different demographics who engage with tweets differently.

Across Tables 3 through 5, there is a moderate correlation between the tweet frequencies of politically right-wing individuals and both real and fake tweets, the correlations are not supported by the low KRCC engagement metrics. Furthermore, the large p-values indicate that the null hypotheses were failed to be rejected, suggesting that politically right-wing individuals' tweets and engagement have little influence on the spread of real and fake tweets. Similarly, politically right-leaning news sources are shown to be moderately correlated with both real and fake tweets in terms of frequency. However, neither of these relationships are statistically significant, thereby requiring further investigation. Furthermore, there is an overwhelmingly low correlation with KRCC values below 0.05 for either group between both engagement metrics and tweet types.

4.2 Sentiment Frequency

Overall, sentiment is not a driving factor in the spread of (mis)information. While real tweets were found to have significant correlations to sentiment frequency with all groups but health organizations and celebrities, this does not align with the findings from section 4.1, suggesting that sentiment does not play a major role in the dissemination of health information. On the other hand, fake tweets have consistently low insignificant sentiment correlations for most groups. Looking at engagement, this lack of influence is further reflected in the abundance of high p-values.

Table 6 shows KRCC values of 0.4086 and 0.2843 for the frequencies of positive and negative tweets, respectively, between political right-wing individuals and real tweets. Both are significant, indicating high and moderate correlations. All other sentiment correlations were insignificant. However, as discussed in section 4.1, political right-wing individuals have little influence on the spread of real and fake tweets. Therefore, while real tweets emote similar moods, it may not be in response (or related) to tweets from political right-wing individuals themselves, but perhaps to similar events to which the individuals are reacting.

Table 6 also shows KRCC values of 0.3506 and 0.2879 for the frequencies of positive and negative tweets, respectively, between political left-wing individuals and real tweets. Both are also significant with p-values of 0.0089 and 0.0330, indicating high and moderate correlation. Similar to right-wing individuals, all other correlations were insignificant. Along with section 4.1, this sentimental significance suggests that those who spread real tweets respond accordingly to both positive and negative sentiments from political left-wing individuals. Considering the positive correlation with frequency and engagement, this match in sentiment provides support for the contribution to the spread of information from political left-wing individuals.

There is a low correlation with regards to sentiment frequency across tweet frequency, average likes, and retweets of health organizations with both real and fake tweets. This low correlation suggests that sentiment is not a factor that drives engagement. Similarly, there are low, insignificant correlations between the sentiment frequency of celebrities and real and fake tweets. With regards to engagement, the only significant mode is for neutral tweets. As discussed in section 4.1, just as celebrities don't contribute to the spread of (mis)information, the lack of sentimental significance also suggests a mismatch in the type of information released between celebrities, and real and fake tweets.

Positive and negative tweets from both right and left news outlets have significant, moderate to strong positive correlations with real tweets. In section 4.1, real tweets show moderate significant correlations in frequency and retweets with left news outlets. This suggests that information from left news outlets are more likely to have been passed on through real tweets, harmonizing in sentiment. For right news outlets, considering low engagement and lack of tweet frequency overlap, the sentimental correlation does not necessarily indicate that the same information is spread, but rather a possibility that the two groups are reacting to similar events. With fake tweets, all correlations were low, with high p-values suggesting a high probability of having no correlation. The observations regarding fake tweets align with the findings from section 4.1, where low correlations were observed for both news sources.

Finally, when looking at engagement with real and fake tweets themselves, low to moderate correlations were found across the tables. However, these were mostly found insignificant, prompting further investigation.

4.3 Sources of Error

A number of future works may be completed to further this study. Currently, the engagement correlations for likes and retweets are not weighted and are treated to have a similar significance as tweet frequency correlations. However, retweets often contribute more to the spread of information compared to likes. Therefore, for future improvements, retweets can be weighted exponentially to mimic the distribution of info through networks while likes can be weighted as a constant. Furthermore, the engagement data used is biased towards individuals who are active on social media and does not account for those who have been affected by the information but do not like or retweet. However, as the focus on the study is the spread of misinformation on social media, it makes sense to not account for the silent users.

There are other limitations to this study, a key source of error being the periods of time each open-source dataset used, affecting the frequency of real and fake tweets. The datasets each focused on a different time frame from which tweets were collected. To counter this, the analysis interval was cut off to between March 1^{st} and May 25^{th}, which had a large number of tweets for each dataset. To ensure each dataset contained the same number of points and to reduce noise, the data was binned every three days. Additionally, there is an uneven number of tweets in each dataset. This is not thought to impact the correlations between each timeline greatly, as the relative fluctuations in frequency should still be correlated.

Furthermore, the classification method of the real and fake tweets differ as they were obtained from different open-source datasets. Some authors used machine learning models, while others relied on fact-checking websites. Additionally, different authors may have had different interpretations of what was considered "fake" and "real". In particular, tweets that fell into a grey-area of "partially" true or false may not be equally classified across different datasets.

What's more, the tweets of influential nodes used were not originally filtered when scrapped by Twint. During data pre-processing, related tweets were filtered on the basis of whether they contained particular keywords. However, it is possible that unrelated tweets may have been included as this filtration method relied on substring searching (e.g. rona is a substring of other common, non-COVID related words). Similarly, related tweets may have been missed as some COVID-related terms were too vague to effectively filter out non-relevant tweets (e.g. while 'vaccine' retrieves tweets related to COVID-19, it may also retrieve tweets related to other illnesses such as the common flu).

Regarding the sentiment analysis, a couple of limitations are noted. This study uses a pre-trained sentiment analysis model from the Python library, Textblob. As the model was not trained to analyze tweets specifically, it is possible that twitter-specific language and slang would not be scored accordingly. Furthermore, the current path of exploration is narrow; while sentiment correlations with tweet frequency may be an indicator that the same information is shared, it does not give insight into how this information is perceived. For example, real information cast in a negative light (e.g. out of disagreement) may affect how their audience perceives and passes on that information (interpreted as real or fake), however, such analysis is difficult to conclude based on this correlation alone. Further work into the sentimental leanings for each group may provide this insight, as well as subjectivity. Finally, the three sentiment modes of focus are broad. For example, 'negative' does not distinguish between sadness and fear, both of which represent very different motivations in the sharing of information.

We also recognize that a low correlation with a high p-value cannot be used to reject the alternate hypothesis that the datasets are correlated; rather, it suggests that it is likely that there is no correlation between the datasets. Further comparisons using a null hypothesis where the KRCC is 1 must be done to reject the correlation.

Conclusion

Living in the Information Age, the wonders of quick info and knowledge access can easily be misused for harmful intents. In difficult times like the COVID-19 pandemic, it is nevertheless more important to be able to learn and identify ways that can reduce the spread of misinformation.

This study examines how specific categories of influential social nodes correlate to the spread of misinformation to find potential ways to limit the propagation of fake news and its harm. The current research and analysis shows that as expected, health organizations strongly correlate to the spread of real news and poorly correlates to the spread of misinformation. Therefore, by encouraging users to follow official medical accounts, this would reduce the exposure and spread of misinformation. Furthermore, the current methods find that politically left individuals help spread more true news than misinformation, and that celebrities drive the spread of true and fake news.

Surprisingly, other categories of influencers have conflicting results or moderate to low correlations to both the spread of true and fake tweets. This could mean that they do not play a significant role in the spread of misinformation. On the other hand, these results can also be due to the sources of errors discussed previously and the use of KRCC instead of other correlation coefficients. Therefore, future work with larger sets of consistent data as well as different correlation measurements should be explored to confirm the validity of the produced results. The sentiment analysis also shows that emotions in tweets do not play a role in spreading misinformation. However, this result could change as we modify and develop our sentiment analysis model.

All in all, we advise Twitter users to follow more health organizations to gain knowledge of real COVID-19 news. Celebrity tweets should be retweeted with caution to make sure as their tweets are a driving force for both real and fake news. Although politically left individuals are shown to help spread more true news than fake news, one should still think before blindly believing every news tweet these individuals post, as they have high correlations with both types of news.

Acknowledgements

We would like to thank Mark Ebden for providing us with valuable advice and guidance on calculating statistics and significance testing. Fur-

thermore, we would also like to thank the STEM Fellowship for this opportunity to learn about data science and engage in infodemiology research.

Appendix

Table 8: Sentiment Average Likes Timeline Correlation KRCC and Mann-Kendall P-Values

		KRCC	P-Value
Health Org.	Pos	-0.1634	0.2154
	Neu	-0.0772	0.5602
	Neg	0.01517	0.9100
Political Right	Pos	-0.3500	0.0085
	Neu	-0.1923	0.1478
	Neg	-0.1461	0.2747
Political Left	Pos	0.2211	0.0946
	Neu	0.3150	0.0178
	Neg	-0.0050	0.9700
Celebrities	Pos	0.1500	0.2591
	Neu	0.2234	0.0940
	Neg	0.2226	0.1057
News Right	Pos	0.1516	0.2519
	Neu	0.2645	0.0446
	Neg	0.1473	0.2672
News Left	Pos	0.1724	0.1974
	Neu	0.2906	0.0272
	Neg	0.2069	0.1201
Real	Pos	0.2150	0.1059
	Neu	0.1732	0.1941
	Neg	0.1978	0.1373
Fake	Pos	-0.0546	0.6795
	Neu	0.0818	0.5355
	Neg	0.1096	0.4166

Table 9: Sentiment Average Retweets Timeline Correlation KRCC and Mann-Kendall P-Values

		KRCC	P-Value
Health Org.	Pos	-0.2376	0.0715
	Neu	-0.1519	0.2517
	Neg	-0.0704	0.5979
Political Right	Pos	-0.2700	0.0422
	Neu	-0.1074	0.4189
	Neg	-0.1915	0.1523
Political Left	Pos	0.2410	0.0685
	Neu	0.3050	0.0218
	Neg	-0.0602	0.6516
Celebrities	Pos	0.2000	0.1324
	Neu	0.2635	0.0482
	Neg	0.1964	0.1534
News Right	Pos	0.1913	0.1481
	Neu	0.2744	0.0372
	Neg	0.1773	0.1818
News Left	Pos	0.1330	0.3234
	Neu	0.2709	0.0401
	Neg	0.1773	0.1846
Real	Pos	0.2650	0.0463
	Neu	0.1079	0.4184
	Neg	0.2128	0.1098
Fake	Pos	0.0893	0.4990
	Neu	0.2404	0.0686
	Neg	0.2065	0.1260

References

[1] More than eight-in-ten americans get news from digital devices. (Accessed on 05/30/2021).

[2] Milos Djordjevic. How many people get their news from social media in 2021?, February 2021.

[3] Majumder S. Chaudhury A., Basuchowdhuri P. Spread of information in a social network using influential nodes. *Lecture Notes in Computer Science*, 7302:121–132, 2012.

[4] Covid-19: 5g broadband conspiracy 'a hoax with no technical basis', un telecoms agency | | un news. (Accessed on 05/30/2021).

[5] World Health Organizaiton. Coronavirus disease (covid-19) advice for the public: Mythbusters.

[6] World Health Organizaiton. Novel

coronavirus(2019-ncov) situation report - 13.

[7] Cole Nemeth. How the twitter algorithm works in 2020. https://sproutsocial.com/insights/twitter-algorithm/. (Accessed on 06/28/2021).

[8] Yonghun Jang, Chang Hyeon Park, and Yeong Seok Seo. Fake news analysis modeling using quote retweet. *Electronics (Switzerland)*, 8(12), 2019.

[9] Ashish Goel, Kamesh Munagala, Aneesh Sharma, and Hongyang Zhang. A note on modeling retweet cascades on Twitter. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9479:119–131, 2015.

[10] Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon. What is twitter, a social network or a news media? In *Proceedings of the 19th International Conference on World Wide Web*, WWW '10, page 591–600, New York, NY, USA, 2010. Association for Computing Machinery.

[11] Yao Lu, Peng Zhang, Yanan Cao, Yue Hu, and Li Guo. On the frequency distribution of retweets. *Procedia Computer Science*, 31:747–753, 2014.

[12] Eytan Bakshy, Jake M Hofman, Duncan J Watts, and Winter A Mason. Everyone ' s an Influencer : Quantifying Influence on Twitter Categories and Subject Descriptors.

[13] Sooraj Divakaran. How to use social network analysis? | digital uncovered. (Accessed on 05/30/2021).

[14] Fink. E Smith R. Understanding the influential people and social structures shaping compliance. *Journal of Social Structure*, 16, 2018.

[15] Salud María Jiménez-Zafra, Antonio José Sáez-Castillo, Antonio Conde-Sánchez, and María Teresa Martín-Valdivia. How do sentiments affect virality on Twitter? *Royal Society Open Science*, 8(4), 2021.

[16] Krishnadas Nanath and Geethu Joy. Leveraging Twitter data to analyze the virality of Covid-19 tweets: a text mining approach. *Behaviour and Information Technology*, 0(0):1–19, 2021.

[17] Candice Lanius, Ryan Weber, and William I. MacKenzie. Use of bot and content flags to limit the spread of misinformation among social networks: a behavior and attitude survey. *Social Network Analysis and Mining*, 11(1):1–15, 2021.

[18] Mary Blankenship. How Misinformation Spreads Through Twitter. pages 1–28, 2020.

[19] Amy Mitchell. *Political polarization & media habits*. Pew Research Center, 2014.

[20] Francesco Poldi Cody Zacharias. An advanced twitter scrapping and intelligence tool.

[21] Matthew Honnibal Roman Yankovsky David Karesh Evan Dempsey Wesley Childs Jeff Schnurr Adel Qalieh Lage Ragnarsson Jonathon Coe Adrián López Calvo Nitish Kulshrestha Jhon Eslava Tyler James Harden Jeff Kolb Daniel Ong Jamie Moschella Roman Korolev Steven Loria, Pete Keen and Ram Rachum. Textblob: Simplified text processing — textblob 0.16.0 documentation. (Accessed on 06/16/2021).

[22] The pandas development team. pandas-dev/pandas: Pandas, Feb 2020.

[23] J. D. Hunter. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, 9(3):90–95, 2007.

[24] Charles R. Harris, K. Jarrod Millman, Stéfan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, and Travis E. Oliphant. Array programming with NumPy. *Nature*, 585(7825):357–362, September 2020.

[25] *Kendall Rank Correlation Coefficient*, pages 278–281. Springer New York, New York, NY, 2008.

[26] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman,

Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020.

[27] Dennis Helsel. *Statistics for Censored Environmental Data Using Minitab® and R.* John Wiley Sons, 2011.

[28] Debanjana Kar, Mohit Bhardwaj, Suranjana Samanta, and Amar Prakash Azad. No rumours please! a multi-indic-lingual approach for covid fake-tweet detection, 2020.

[29] William Scott Paka, Rachit Bansal, Abhay Kaushik, Shubhashis Sengupta, and Tanmoy Chakraborty. Cross-sean: A cross-stitch semi-supervised neural attention model for covid-19 fake news detection, 2021. (Accessed on 05/31/2021).

[30] Mohamed K. Elhadad, Kin Fun Li, and Fayez Gebali. Covid-19-fakes: A twitter (arabic/english) dataset for detecting misleading information on covid-19. In Leonard Barolli, Kin Fun Li, and Hiroyoshi Miwa, editors, *Advances in Intelligent Networking and Collaborative Systems*, pages 256–268, Cham, 2021. Springer International Publishing.

[31] Gautam Kishore Shahi, Anne Dirkson, and Tim A Majchrzak. An exploratory study of covid-19 misinformation on twitter. *Online Social Networks and Media*, 22:100104, 2021.

[32] Shahan Ali Memon and Kathleen M. Carley. Characterizing covid-19 misinformation communities using a novel twitter dataset, 2020. (Accessed on 05/31/2021).

[33] StatsDirect. Kendall's rank correlation.

[34] Mostafavi. A Chao R., Jiang. Y. The role of local influential users in spread of situational crisis information. *Journal of Computer-Mediated Communication*, 2021:108–127, 2018.

[35] Lee Rainie, Aaron Smith, Kay Lehman Schlozman, Henry Brady, and Sidney Verba. Social media and political engagement, 2012. (Accessed on 05/31/2021).

[36] Eva Lahuerta-Otero, Rebeca Cordero-Gutiérrez, and Fernando De la Prieta-Pintado. Retweet or like? that is the question. *Online Information Review*, 2018.

[37] Ruben Geert van den Berg. Kendall's tau - simple introduction. (Accessed on 05/31/2021).