Twitter Network Analysis of Lockdowns in Light of Covid-19

Jeffrey Kwan

Table of Contents

Introduction

In this day and age, social media is a powerful tool for communication. While social media could be used for sharing moments from one's own life, it is also used for transmitting information and expressing one's opinion.

Lockdowns refer to mass stay-at-home orders. This technique is effective in isolating individuals and stopping the spread of infectious diseases. In light of the Covid-19 pandemic, many countries have resorted to nationwide lockdowns, as hospitals became overwhelmed with the exponentially increasing cases. By the end of March, over 100 countries had initiated some kind of lockdown ("Coronavirus: The world in lockdown", 2020). These measures sparked a plethora of discussion on popular social media platforms, such as Facebook and Twitter. Many were critical of the draconian measures taken by countries like China (and some others) (Malik, 2020), but public opinion seemed to have shifted in support of lockdowns after countries with no lockdown policies like the UK and Sweden saw a spike in cases.

What exactly causes these fluctuations? How do people's opinions change? To what extent do major events and announcements shift public opinion? How do people with different opinions interact with each other? This paper attempts to answer these questions.

Volume analysis was performed on two sets of historical Twitter data and network analysis was performed on a set of real-time data. The results indicate that real-life events have an effect on the volume of discussion around the topic of lockdowns, in light of Covid-19 and the overall opinion of the community. Network analysis confirms Himelboim et al.'s (2016) results that users with polar opinions on lockdowns tend to separate into two factions, while neutral users are not excluded by either faction.

Literature Review

Twitter analysis has been used extensively for marketing analysis and for academic research. On the academic side, many studies have been conducted relating to the structure of Twitter networks and the spread of information on the platform, be it on “Occupy Wall Street” (Tremayne, 2014) or on warnings for natural disasters (Chatfield, 2012).

Himelboim et al. (2016) studied the network structure surrounding controversial topics in the 2012 Presidential Election. Regarding the recent coronavirus pandemic, Ahmed et al. (2020) investigated the 5G conspiracy theory by identifying community structure and key users to see what public health authorities can do to mitigate the spread of disinformation. Himelboim et al.’s findings have shown that people with the same political views tend to form tight communities and rarely interact with people of differing or conflicting views.

A common method used to identify the opinions of text in a systematic fashion is sentiment analysis. Many sentiment analysis algorithms use Natural Language Processing (NLP) and machine learning to identify sentiment based on individual word connotations, and also based on context. Specifically, an open-source sentiment analysis tool called Vader (Valence Aware Dictionary and Sentiment Reasoner) was used in this investigation. Vader uses a lexicon and rule-based approach that works especially well for social media (Hutto & Gilbert, 2014) and in the pilot tests.

Model

All coding is done using Jupyter notebook and Visual Studio Code in Python. There are two parts to this analysis: Volume Analysis and Network Analysis. The first part investigates how real-life events influence Twitter activity, and the second part explores the network structure of people talking about lockdowns and the dynamics of opinionated networks.

Volume Analysis

We model the change in Twitter activity as 1) change in volume of tweets and 2) change in people's opinions. As such, we split this section into two parts: General Volume Analysis and Sentiment Analysis. Two key events in the early stages of the pandemic will be analyzed:

1. When Wuhan announced lockdown on January 23rd, 2020 (Regan et. Al, 2020).
2. When Donald Trump signed an $8 billion package aid package on March 6th, 2020 (Hirsch & Breuninger, 2020).

General Volume Analysis

Tweets containing the keywords "lockdown," "quarantine," and "reopen" are gathered using a Python library called twint,[1] from January 15th to Janurary 30th, 2020, and March 3rd to March 9th, 2020. Each tweet is scraped for data about the username, date, the tweet, as well as the number of replies, retweets, and likes. The Twitter activity[2] is graphed per day. The dataset collected for the Wuhan lockdown consists of 111,761 tweets and the dataset collected for Trump's aid package signing consists of 225,714 tweets.

Sentiment Analysis

Sentiment analysis is performed on each tweet using Vader. For the purposes of this investigation, the output is simplified to a trinary of negative, neutral, and positive. A more detailed description can be found in the Appendix A. As a general assumption, pro-lockdown tweets are likely to say positive things about lockdowns and anti-lockdown tweets are likely to

1 https://github.com/twintproject/twint/wiki

2 This refers to the total number of tweets, retweets, and replies.

be negative. Hence, in this paper, positive refers to pro-lockdown, negative refers to anti-lockdown, and neutral represents tweets without a strong opinion.

The total number of tweets per day containing each sentiment is also graphed. While it is easy to discern visually whether there are more positive or negative tweets, it is harder to determine overall how "positive" or "negative" the tweets are. Hence, a separate graph is plotted for the "overall polarity" of tweets each day, by taking the average of the compound scores of the tweets on each day.

Network Analysis

An interaction network is obtained with the nodes as users and the edges as an interaction between users, namely a mention, a reply, or a retweet. The dataset for this section is collected with the keyword "quarantine" using the Python library Tweepy[3], from 7:54pm, June 19, 2020 PST, to 7:15am, June 20, 2020 PST. It is 1.58GB in size and contains 247,584 tweets.

General Network Analysis

Inzaugarat (2019) provides a good approach for generating Twitter interaction networks using the Networkx, so I have adapted her code for this investigation. A brief summary of the workflow can be found in Appendix B. As edges are defined by interactions between users, interactions are obtained from the dataset by taking the user's ID of the post and the ID of the user who was mentioned/replied to/retweeted.

[3] http://docs.tweepy.org/

Network Sentiment Analysis

Sentiment analysis is used to determine the overall stance of each user on lockdowns (positive, negative, neutral) and the color of each node is changed accordingly (discussed further in Appendix C). For visualization and convenience purposes, we define the nodes' sentiments based on their color. Thus, users with generally positive tweets will be referred to as *blue nodes*, those with negative tweets are *red nodes*, those with a neutral stance are *green nodes*, and the users whose sentiments could not be determined will be named *grey nodes*.

The community structure of the different colored nodes is determined. This is done by determining average percentages of different color combinations—how often a red node neighbors[4] a blue node, for example. The values of these percentages determine how much each node color interacts with nodes of the same or other colors. For example, a high red percentage for red nodes suggests that red nodes tend to interact with other red nodes, which translates to a high clustering coefficient, and a low average path length among red nodes.

Results and Analysis[5]

Part 1: Volume Analysis

General Volume Analysis

From Figure 1, a noticeable jump in volume can be seen from January 21st to 23rd. This coincides with the announcement of the lockdown in Wuhan. However, Figure 2 shows a more gradual increase compared to Figure 1 and approximates the shape of a polynomial graph. It is difficult to compare the two, as the data collected for this event only spans 1 week due to time constraints, but there seem to be differing degrees of sharp transitions between low and

[4] The Networkx documentation defines a neighbor node q of node p as being able to form an edge (directed or undirected depending on the graph) from p to q. For more: https://networkx.github.io/documentation/stable/reference/classes/generated/networkx.DiGraph.neighbors.html

[5] All Tables and Figures can be found in Appendix C.

high volumes of tweets. For example, Figure 1 would appear to have a “sharper” transition than figure 2. Is there a correlation between this and the “width of transitions” used in epidemiology? This merits an interesting investigation for future research.

Adding the retweet and reply numbers to the equation (Fig. 3) seem to follow the same general trend as Figure 1, although there is an unexplained dip from the 26th to the 30th. Similarly, Figure 4 follows the same general trend as Figure 2. From Figure 1 through 4, it can be reasoned that tweet activity is, to some extent, proportional to the total Twitter activity. Therefore, if there are time constraints, it may be suitable to only gather tweet data and to ignore the retweets and the replies.

Sentiment Analysis

Figures 5 and 6 show an overwhelming number of neutral tweets, which could be attributed to the many news article tweets found in the pilot tests. It could also be attributed to Vader’s misclassification bias which will be further discussed in the Limitations section.

In both Figures, the three sentiment plots approximately follow the shapes of Figures 1 and 2. However, in Figure 6, the number of negative tweets were consistently higher than the number of positive tweets, while Figure 5 shows a spike in negative tweets following the Wuhan lockdown.

While one could mistake the sharp increase in total tweets for the sharp increase of negative tweets in Figure 5, Figure 7 shows that the overall sentiment did shift from positive to negative on January 22nd. Figure 8 shows that not only did the total number of negative tweets increases (Fig. 6), the overall negativity increased six-fold on March 7 compared to the day before, which is significant considering the number of neutral tweets that would have diluted the results. Once again, this correlates to Trump’s signing of the aid package. There is one complication here. Firstly, Trump signed the aid package on March 6, not 7, but the delay can be explained by news outlets giving information at a delayed rate.

As a conclusion, the two major announcements in the coronavirus timeline seem to correlate to a shift in public sentiment. Comparing the two case studies, since the coronavirus was not a very hot topic in January, the Wuhan announcement shows that people had some sort of panic reaction as they took to Twitter: the number of tweets increased and the overall sentiment dipped into the negatives, while in March, when the pandemic was already brewing for some time, people seemed to be more negative about the outlook of the situation, which became even more negative with the Trump announcement.

Part 2: Network Analysis

For the purposes of this paper, we define *all graph* as the graph with edges of all three types of interactions (replies, retweets, mentions) and the nodes associated with the edges. We also define *reply graph*, *retweets graph*, and *mentions graph* as the network whose edges are solely reply, retweet, and mention interactions respectively, and the nodes are the users associated with the edges.

General Network Analysis

Tables 1, 2, and 3 (Appendix D) show the general network properties of each graph. Broadly, this network is one of singular interactions, where users mostly retweet or sometimes comment on a particular post and do not engage in much meaningful conversation to support or refute perspectives. This could also be due to the nature of Twitter, as retweets tend to be one-directional and do not invoke further interaction. A more detailed analysis can be found in Appendix E.

Considering the small size of the giant component[6] and content of tweets of the interactions that lie outside of the giant component, it would appear that interactions that happen outside of the giant component are irrelevant to opinions on lockdowns and the focus of this paper.

Network Sentiment Analysis

According to Table 4, neutral users seem to be in the vast majority. This seems to be a constant in all the datasets that were used in this investigation.

From tables 5 and 6, it can be observed that red nodes tend to interact with other red nodes and green nodes. Similarly, blue nodes tend to interact with other blue nodes and green nodes. It can also be seen that red nodes rarely, if ever, interact with blue nodes, and vice versa. These observations correspond with Himelboim et al.'s (2016) results: people of similar (political) opinions tend to form isolated communities. It cannot be concluded, though, that red and blue nodes form *tight* communities, as calculations concerning clustering could not be made in this paper, due to time and skill constraints.

One seeming contradiction is that the dominant color that red and blue nodes interact with is neither red nor blue, but rather the green nodes. This is a novel observation as Himelboim et al. (2016) have not considered the neutral opinion in their paper. A plausible explanation is through the consideration of the relative numbers of each color of nodes. As can been seen from Table 4, green nodes far outnumber red and blue nodes. Consequently, it becomes far more likely for red and blue nodes to interact with green nodes, even if this is only a small percentage of green nodes. This is further reinforced by the green column in Table 6, which shows extremely low percentages of blue and red neighbors. This explanation leads to a

[6] The giant connected component is the largest connected subgraph of a network.

secondary non-trivial conclusion: positive and negative users do not repel neutral users. Considering for a second the sentiments that each color represents, this may be due to neutral users tweeting out news articles, and the red and blue users retweeting or replying to these articles to help reinforce their case. Through sampling some of the tweets between red and green users and between blue and green users (Appendix F), the conjecture appears to hold true for this network.

One of the cases that Kleinberg (2010) considers as a structurally balanced network is one in which two factions strongly interact with members of the same faction and rarely with the other faction. Looking at the data, this Twitter interaction network appears to be structurally balanced. However, drawing a concrete conclusion demands more careful analysis, as we have not considered the nature of interactions between the reds and blues.

Bringing the findings from the General Network Analysis section back into the discussion, it seems that—at least on Twitter—people with strong opinions tend to surround themselves with others of similar opinion. This only serves to strengthen one’s own stance while completely disregarding the perspectives of others. They also draw resources from neutral sources.

Limitations

There are numerous limitations to this investigation, the three main ones being time, processing power, and expertise/knowledge. Without these limitations, a few significant and effective improvements could have been made. A brief description of each limitation will be listed here. These are discussed further in Appendix G.

1. Gathering data for the entire 6+ month duration of the pandemic.
2. Performing intent analysis to determine pro- and anti- lockdown sentiment instead of brute forcing with sentiment analysis.

3. Utilizing natural language processing and machine learning to gather more relevant tweets instead of a keyword search.
4. Utilizing a supervised machine learning approach to patch Vader's pre-trained model for more accurate classification of tweets.

Conclusion

As a resolution to the research objectives, a few conclusions can be drawn.

1. Intuitively, major real-life events cause rises in Twitter activity, and the analysis performed in this investigation was unable to refute this hypothesis. The rate at which Twitter activity increases seems to follow a "width of transition" of sorts.
2. The data shows that public opinion seems to shift from positive to negative or from negative to even more negative before and after major events. This evidence strongly suggests that the events do play a role in shifting public opinion. In the context of Covid-19, major announcements seem to spread negativity.
3. Interaction communities on Twitter on the topic of lockdowns are sparse and one-directional. People of strong opinions surround themselves with like-minded people while ignoring those with other strong opinions. They acquire material from neutral or like-opinionated. This creates a community of opinions that ignore any form of counterevidence.

Throughout the project, many interesting results were discovered, which opened up a multitude of directions for future research. Three in particular stand out for closer consideration. These include performing user analysis on the dataset in Network Analysis to determine how factions are formed—by influential users or organically. Most intriguing, however, will be investigating the "width of transition" phenomenon observed in Volume Analysis, and diving a bit deeper into the idea of structural balance (Kleinberg, 2010) and how the network in Network Analysis relates to this concept.

References

Ahmed, W., Vidal-Alaball, J., Downing, J., & Seguí, F. L. (2020). Covid-19 and the 5g conspiracy theory: Social network analysis of Twitter data. Journal of Medical Internet Research, 22(5), e19458.

Chatfield, A., & Brajawidagda, U. (2012). Twitter tsunami early warning network: a social network analysis of Twitter information flows.

Coronavirus: The world in lockdown in maps and charts. (2020, April 07). Retrieved from https://www.bbc.com/news/world-52103747

Easley, D., & Kleinberg, J. (2010). Networks, crowds, and markets (Vol. 8). Cambridge: Cambridge university press.

Himelboim, I., Sweetser, K. D., Tinkham, S. F., Cameron, K., Danelo, M., & West, K. (2016). Valence-based homophily on Twitter: Network analysis of emotions and political talk in the 2012 presidential election. New media & society, 18(7), 1382-1400.

Hirsch, L., & Breuninger, K. (2020, March 06). Trump signs $8.3 billion emergency coronavirus spending package. Retrieved from https://www.cnbc.com/2020/03/06/trump-signs-8point3-billion-emergency-coronavirus-spending-package.html

Hutto, C.J. & Gilbert, E.E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. Eighth International Conference on Weblogs and Social Media (ICWSM-14). Ann Arbor, MI, June 2014.

Inzaugarat, E. (2019, May 08). Visualizing Twitter interactions with NetworkX. Retrieved from https://medium.com/future-vision/visualizing-Twitter-interactions-with-networkx-a391da239af5

Malik, K. (2020, February 02). Locked-down Wuhan and why we always overplay the threat of the new | Kenan Malik. Retrieved from https://www.theguardian.com/commentisfree/2020/feb/02/coronavirus-wuhan-lockdown-breeding-ground-fear-stigma

Regan, H., Griffiths, J., & Culver, D. (2020, January 24). Wuhan coronavirus virus spreads as China scraps New Year celebrations. Retrieved from https://www.cnn.com/2020/01/23/china/wuhan-coronavirus-update-intl-hnk/index.html

Tremayne, M. (2014). Anatomy of protest in the digital era: A network analysis of Twitter and Occupy Wall Street. Social Movement Studies, 13(1), 110-126.

Appendix A

Vader Scoring Criteria

Normally, Vader outputs a compound score that determines the polarity of the tweet from -1 to 1, but for the purposes of this investigation, the output is simplified to a trinary of -1 for negative, 0 for neutral, and 1 for positive. For example, “Staying home helps flatten the curve!” would be considered a 1; “The total cases of Covid-19 has risen” would be a 0; and “I hate this lockdown!” would be -1. Thresholds are used to simplify the compound score. In this section, >=0.5 was considered positive, <= -0.3 was considered negative, and anything in between was neutral. The positive threshold was raised because many neutral tweets were systematically misclassified as positive.

Appendix B

Workflow for Interaction Network Generation

- A directed graph is generated such that isolated nodes are removed.
- Basic network properties are computed and network properties for community structure (clustering, average path length, etc.) are computed by taking a 1000 sample of nodes due to efficiency reasons and hardware constraints.
- Steps 8 – 12 are repeated with each of the replies, retweets, and mentions to investigate the properties of each type of interaction.

Appendix C

Sentiment Scoring Criteria for the Interaction Network

Before diving into the details, the structure of the data sent by the Twitter API must first be explained, in order to understand the following. As mentioned earlier, Twitter stores every tweet and the information associated with it in nested json objects. Among them, there are items called “in_reply_to_user_id,” “retweeted_id,” “user_mentions_id,” and “user_id”. If the tweet were a retweet, the “retweeted_id” cell would contain the user ID of the original poster of the tweet. If the tweet were a reply, the “in_reply_to_user_id” would contain the user ID of the tweeter whose tweet was replied to, and so on.

A Vader sentiment score is given to the text and attributed to the user’s ID and the retweeted ID, because a tweet and a retweet have the same content. To obtain every unique ID from the dataset, if “in_reply_to_user_id” and “user_mentions_id” are not empty, a null value is given to them because there is no text which can be used to determine their sentiments.

The overall sentiment of a user is determined by averaging the sentiment scores of every tweet associated with the user. It is then simplified to the trinary seen in the Sentiment Analysis section with the following criteria: >= 0.3 for positive, <= -0.3 for negative, and anything in between as neutral.

Appendix D

Tables and Figures

Figure 1

Number of tweets about lockdowns per day from Jan 15 – 30

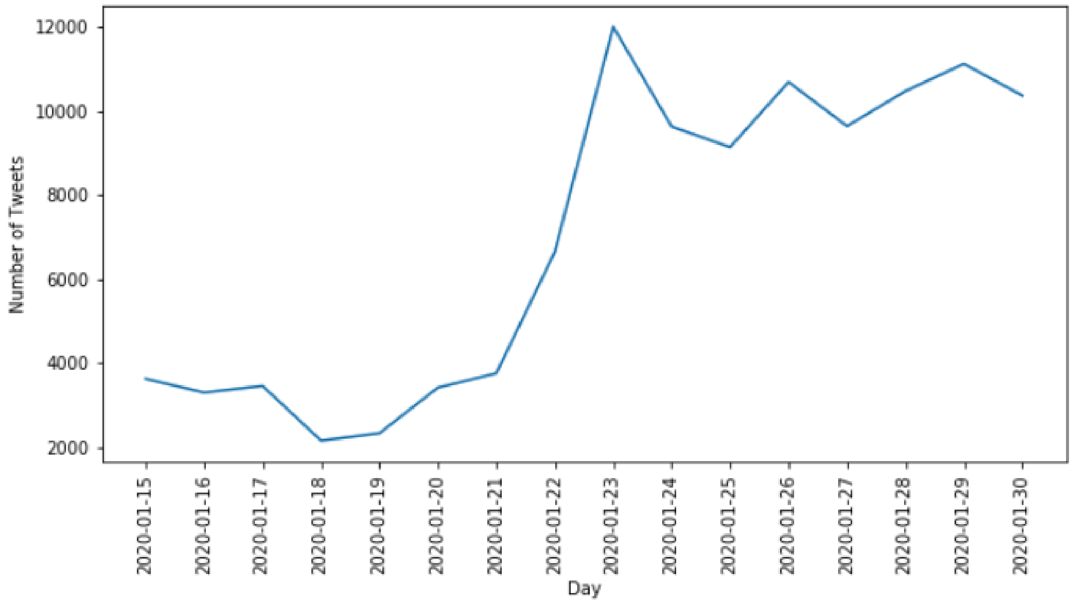

Figure 2

Number of tweets about lockdowns per day from March 3rd – 9th

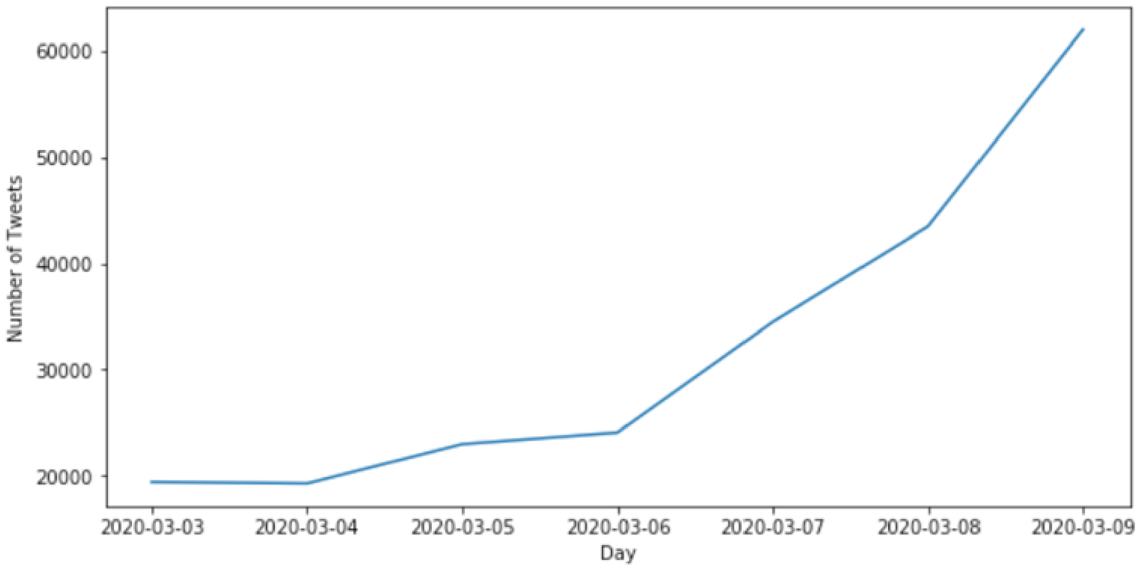

Figure 3

Total Twitter activity about lockdowns from January 15 - 30

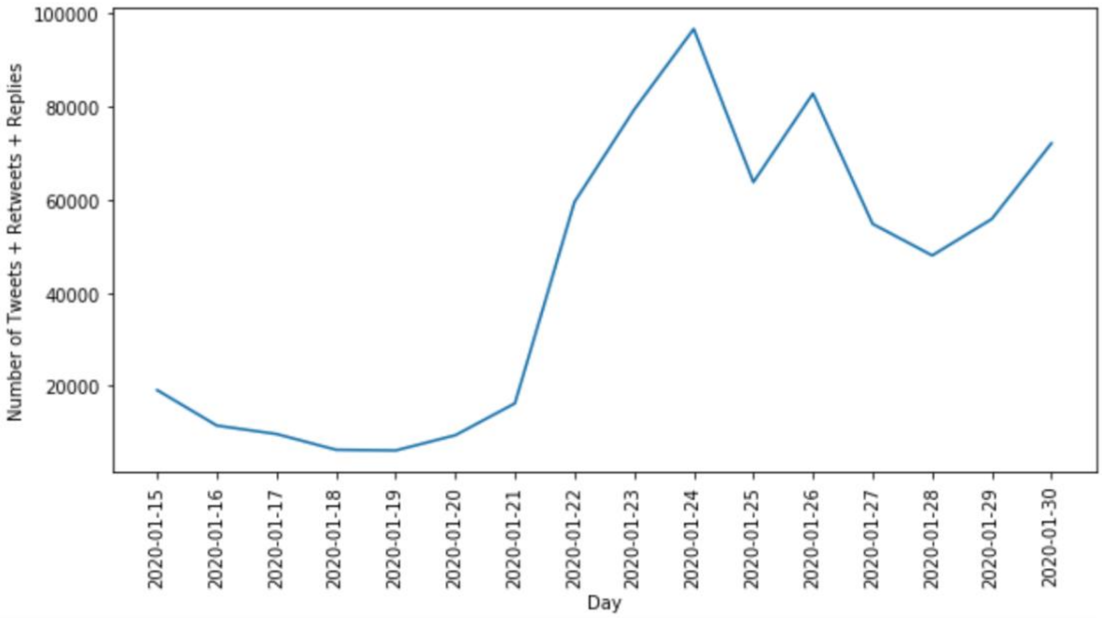

Figure 4

Total Twitter activity about lockdowns from March 3 – 9

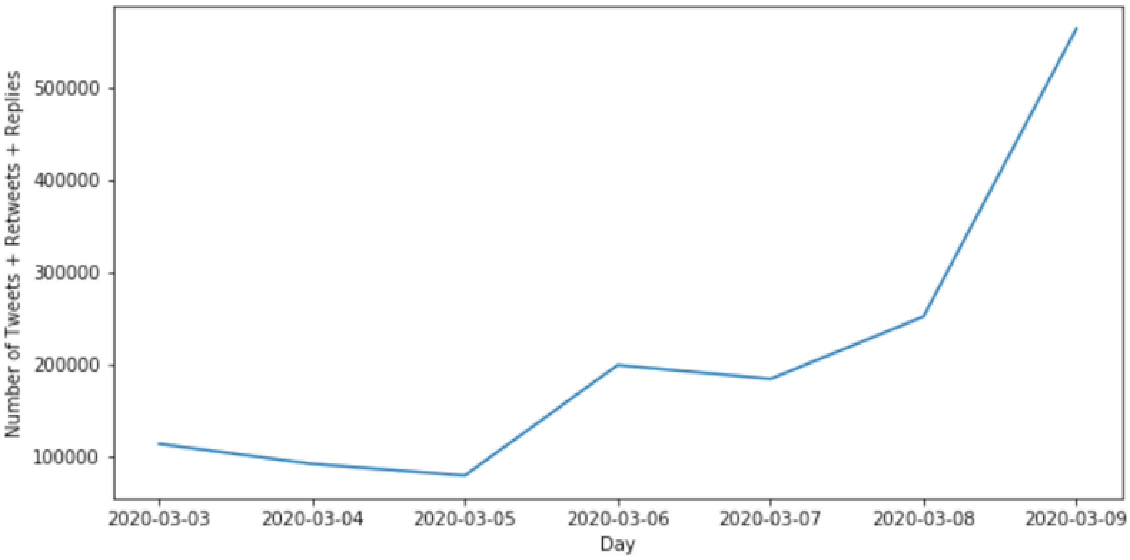

Figure 5

Number of tweets of each respective sentiment from Jan 15 – 30

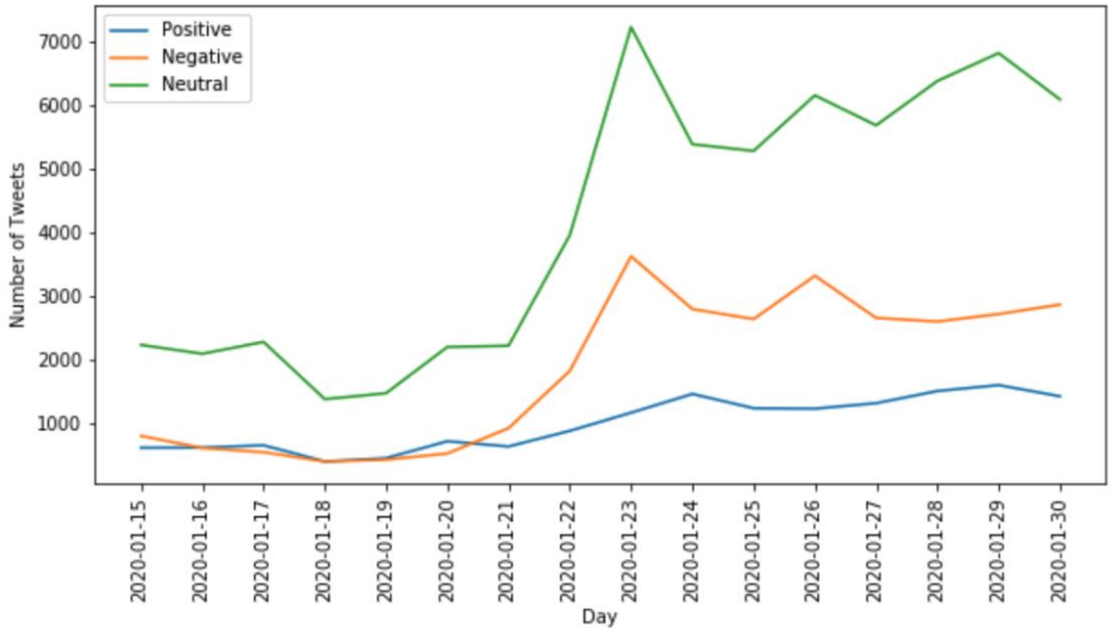

Figure 6

Number of tweets of each sentiment from March 3 – 9

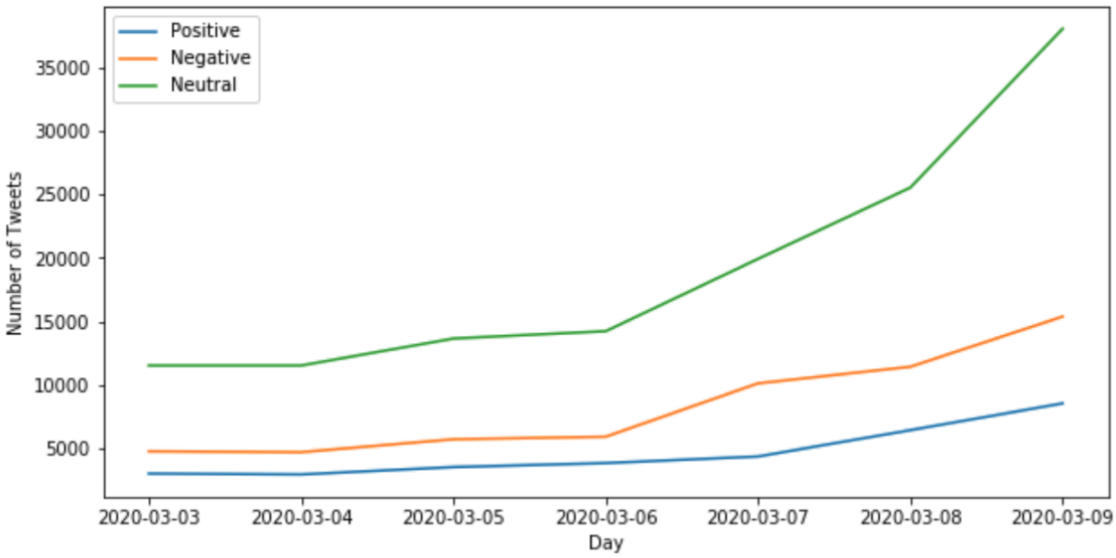

Figure 7

Overall sentiment of tweets about lockdowns from Jan 15-30 on a scale of -1 to 1.

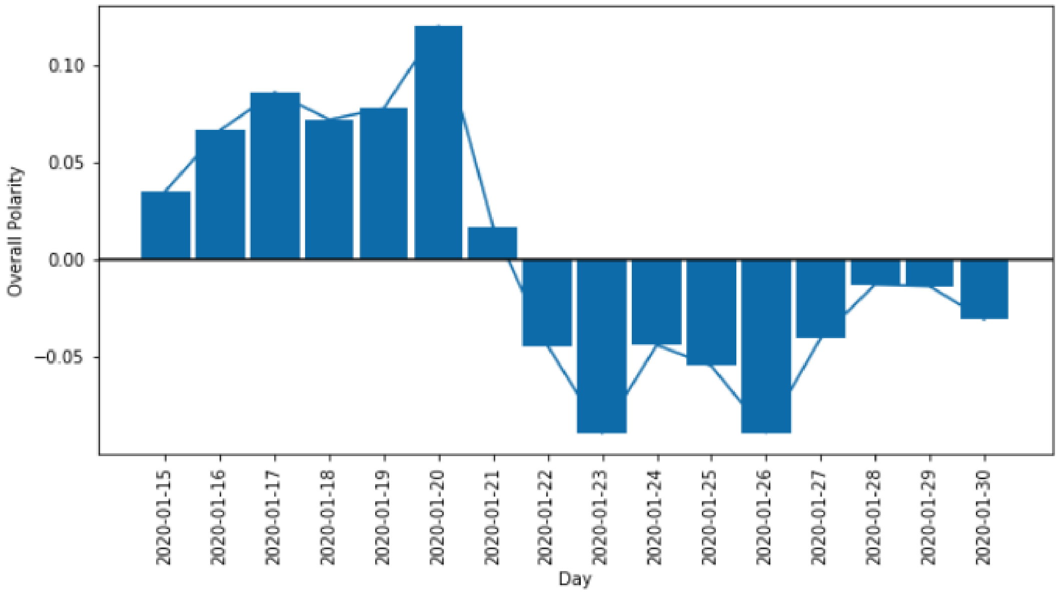

Figure 8

Overall sentiment of tweets about lockdowns from Mar 3-9 on a scale of -1 to 1.

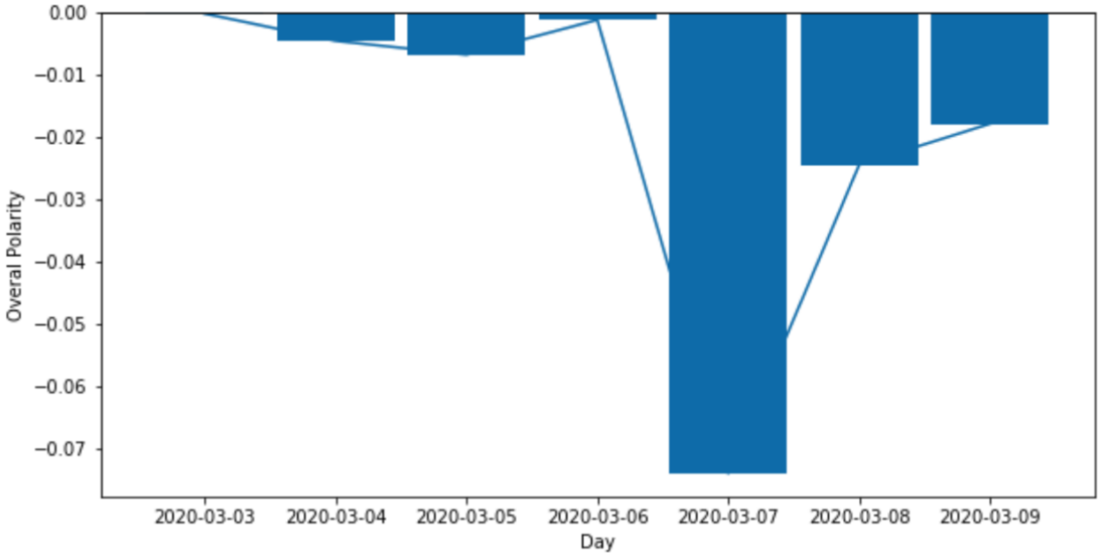

Table 1

Network Size properties

	All graph	Replies graph	Retweets graph	Mentions graph
# nodes	83374	14959	62315	75562
# edges	81238	8768	62801	72017
# connected components	11238	6209	5667	10365
# nodes in GCC[7]	41336	548	38064	40590
# edges in GCC[8]	48435	562	43559	46439

Table 2

Degree Properties

	All graph	Replies graph	Retweets graph	Mentions graph
Max Degree	11185	237	11185	11185
Min Degree	0	1	1	1
Max In Degree	11185	237	11185	11185
Min In Degree	0	0	0	0
Max Out Degree	64	30	64	64
Min Out Degree	0	0	0	0
Avg Degree	1.9	1.2	2.0	1.9

7 Giant Connected Component.
8 Note that this value was calculated from the undirected version of each graph due to software limitations, so some bidirectional directed edges from the original graph may have merged. Therefore, the value displayed is likely a bit lower than the actual number.

Most frequent degree	1	1	1	1
Avg In Degree	1.0	0.6	1.0	1.0
Most frequent In Degree	0	0	0	0
Avg Out Degree	1.0	0.6	1.0	1.0
Most frequent Out degree	1	1	1	1

Table 3

Community structure properties

	All graph	Replies graph	Retweets graph	Mentions graph
Average clustering (±3%)	0.0049	0.000	0.0026	0.0037
Diameter (±3%)	31	19	26	37
APL (±3%)	9.13	7.79	8.90	9.228
Avg. # nodes in CC[9]	3.74	2.32	4.28	2.67
Avg. # edges in CC	2.89	1.32	3.38	1.67

Table 4

Number of nodes in each color

Red	Blue	Green	Grey
16165	20575	47014	12591

9 Note that the giant connected component was excluded from these calculations.

Table 5

Proportion of neighboring nodes of each color in the undirected all graph. The average percentage of nodes that are the same color as the node examined is highlighted in bold lettering.

% of neighbors are:	Red	Blue	Green	Grey[10]
Red	**40.15%**	0.20%	0.90%	14.68%
Blue	0.129	**37.54%**	1.28%	29.01%
Green	54.90%	55.18%	**91.88%**	56.31%
Grey	4.66%	7.08%	5.94%	0.00%

Table 6

Proportion of neighboring nodes of each color in the directed all graph.

% of neighbors are:	Red	Blue	Green	Grey
Red	**35.55%**	0.15%	0.78%	0.00%
Blue	0.22%	**28.81%**	1.04%	0.00%
Green	55.63%	56.68%	**87.49%**	0.00%
Grey	4.82%	7.30%	6.18%	0.00%

10 While grey nodes have to be accounted for in order to preserve all the nodes in the interaction network, they are irrelevant to the focus of this section.

Appendix E

Extended General Network Analysis

Retweets and mentions dominate the interactions in this dataset. As retweets involve broadcasting someone else's tweet to one's own audience, the network appears to be a broadcast network that distributes tweets to a wider audience. Judging by the nature of tweets seen in the Volume Analysis section, many of these retweets would appear to be for news articles.

While there does exist a node with a rather high in-degree, most of the nodes in the graph seem to have an average degree of 0 or 1. Considering the low degree, as well as the clustering, high diameter and average path length, a Twitter interaction network does not tend to form tightly knit communities. Coupled with the fact that the number of nodes and edges seem to be roughly equal in all the graphs except for in the reply graph, this translates to a network that is not made of small discussion communities, but rather a network of singular interactions.

Appendix F

Sampling of Red-Green and Blue-Green Tweets

Here are the first five tweets from red to green users.

RT @ndtv: Watch | "People who have mild symptoms or are asymptomatic can recover within the comfort of their homes": Atishi, MLA, A
AP on ce…

RT @Chaiti: Just doesn't make sense. World over we have seen how home isolation is an option for mild to moderate symptoms. Hospital
s can n…

RT @ANI: #WATCH - Centre's decision of making 5-day institutional quarantine compulsory for #COVID19 patients in Delhi is wrong. I am
a cor…

RT @bsindia: Centre's decision of making 5-day institutional quarantine compulsory for Covid-19 patients in Delhi is wrong: @AtishiAAP

RT @sardesairajdeep: In other news.. Delhi LG passes order of mandatory 5 day govt quarantine for those testing positive for corona. NO
hom…

4 out of 5 tweets can be said to be news headlines or descriptions of them.

Here are the first five tweets from blue to green users.

RT @drkerrynphelps: This is not over #COVID19
Coronavirus Victoria: Hundreds of Melbourne hotel workers in quarantine as security guards te…

RT @myworld2121: Quarantine results … !!

(whiskyhuskydog | Ig) https://t.co/5AS9Tafqnr

RT @interaksyon: As the nation eases into a modified community quarantine status, the logistics industry—including transportation, stora
ge,…

RT @interaksyon: After mañanita, Filipino online users are now criticizing the "despedida party" involving some BFP Region VI personnel t
h…

RT @ANI: Delhi Chief Minister Arvind Kejriwal at today's SDMA meeting said there is already a shortage of healthcare staff, how will it b
e…

Similarly, 3 or 4 out of the 5 tweets are related to the news.

Appendix G

Limitations (continued)

1. Due to time and processing constraints, only data from the initial stages of the pandemic were able to be gathered. Additionally, only one week's worth of data was collected for the Trump aid package event, which made it difficult to draw reliable conclusions from the data. If it were not for these constraints, gathering data for the entire 6+ month duration of the pandemic would have been optimal, as conclusions would have been drawn based on the data, instead of checking to see if the data fits predictions.
2. A major limitation of this project hinges on the assumption that a positive sentiment refers to pro-lockdown and a negative sentiment refers to anti-lockdown. While this is true for the most part, positive and negative do not directly mesh with the two categories. For example, the sentence "quarantine is finally over!" would likely be classified as a positive tweet, even though it conveys a neutral, if not anti-lockdown stance. Some spot checks are done in Appendix H. Hence, to mitigate this issue, intent analysis could have been used to determine a more accurate representation of pro-lockdown and anti-lockdown sentiment.
3. Doing a keyword search for tweets is a crude way to gather tweets about a certain topic, as presumably many Twitter users discuss the topic at hand using synonyms or other related words that can never be fully captured in an exhaustive list. Adding the number of retweets and replies to the number of tweets reduced this limitation to some extent, but a more effective solution for this would be to use natural language processing and machine learning to find more relevant words and gather more relevant tweets. Due to time constraints and the limitations of my knowledge in machine learning, however, this was not feasible for this paper.

4. While doing pilot tests, Vader is seen to perform poorly with sarcastic tweets. As a general comment, Vader generally tends to misclassify tweets as neutral more than it would positive or negative. On top of that, “quarantine” as a verb tends to cause Vader’s classification to go awry. Some examples of this can be found in Appendix I. Since the use of “quarantine” as a verb usually has a negative connotation, a supervised machine learning approach could be used to patch Vader’s pre-trained model to more accurately classify tweets of this nature. This problem seemed to only occur with the datasets used in Volume Analysis, so the threshold for positive classification remained at standard levels when performing sentiment analysis on the dataset used in Network Analysis.

Appendix H

Spot Checks for Assumptions

RT [redacted] Safe pair of hands ... everyone Cabinet needs one.
have u ever done this during quarantine? yes/no

RT [redacted] I've met the best people throughout this quarantine like legit made me feel so happy, can't wait to see all of them that'll···

These are sample tweets taken from the positive classification. They do not seem to be pro-lockdown necessarily, but they are not anti-lockdown either. If anything, these tweets show that people are taking quarantine positively, so it could be argued that they support lockdowns.

Have to be somewhere (dressed, hair, makeup) before 9 AM tomorrow for the first time since quarantine and I am sett··· https://t.co/LUKSFov5XK
At the stage of quarantine where she finds my water cup selection annoying😒🙄😥

These are sample neutral tweets. They also do not display strong opinion towards lockdowns are justified remaining where they are.

seriously bro, cases are still rapidly increasing. I feel so disappointed.
Fact: With our current (inadequate) test & trace capabilities, we have nearly no way to control the spread administ··· https://t.co/uhs0hlmEX9
RT [redacted] The Philippines has had the longest lockdown, and it's been ineffective as [redacted] Aside from extending quarantine, WHAT IS O···

These are some of the negative tweets. The tweets convey a strong anti-lockdown stance, so they are justified as being classified as negative.

Appendix I

Examples of Vader's Classifications

Alamak GG, let's quarantine together then not boring 😂 {'neg': 0.108, 'neu': 0.566, 'pos': 0.326, 'compound': 0.5943}

This tweet was correctly identified as positive. Vader demonstrates its strong ability to accurately parse emojis and other social media exclusive symbols and text.

Gosh. Pls quarantine them completely and cancel all in and out flight from China from the time being. {'neg': 0.116, 'neu': 0.817, 'pos': 0.066, 'compound': -0.2449}

This tweet was correctly classified as negative, but was more neutral-leaning than it should have been.

Yeah but honestly there's no reason to quarantine entire 20 million population just over few deaths, numbers are definitely dodgy here. {'neg': 0.17, 'neu': 0.528, 'pos': 0.302, 'compound': 0.6124}

This is undoubtedly a neutral tweet but was classified as quite highly positive.

```
In [144]: sentiment_analyzer_scores('Probably best to quarantine  China. .... just saying. ...  #CoronavirusOutbreak')
Probably best to quarantine  China. .... just saying. ...  #CoronavirusOutbreak {'neg': 0.0, 'neu': 0.682, 'pos': 0.318, 'compound': 0.6369}
```

This tweet is supposed to be neutral, even leaning to the negative side, but was classified as highly positive.

Thx for ur tweets, very useful. Re: quarantine, do u think it has use in reducing velocity of spread around China+world, compared to if no quarantine at all? I.e. is it "horse out of barn, too late to shut door" or "1 horse out of barn, but some still inside, so shut the door"? {'neg': 0.028, 'neu': 0.898, 'pos': 0.074, 'compound': 0.3831}

This is a neutral tweet that was almost classified as positive. There were many other neutral tweets that were classified in the same way which had to be compensated for by raising the threshold for the neutral classification.