SENTIMENT ANALYSIS ON THE TWITTER USAGE OF G7 WORLD LEADERS DURING COVID-19

NEELESH VASNANI

OCTOBER 2021

CAPSTONE DISSERTATION

MASTER OF ANALYTICS PROGRAM

MASSEY UNIVERSITY

TABLE OF CONTENTS

ABSTRACT

During public health emergencies and crisis situations, citizens increasingly rely on world leaders for information. Social media platforms such as Twitter have emerged as powerful communication tools for world leaders to rapidly share information and reassure the public during crises. The information, language, and tone within the leaders' online posts can steer the opinions, behaviors, and mental states of citizens. Given this influence, leaders have a political, professional, and ethical responsibility to provide swift, transparent, and relevant information to citizens. To knowledge, this is the first study to apply sentiment analysis and unsupervised topic modeling using Natural Language Processing (NLP) to explore how the Group of Seven (G7) leaders used Twitter to communicate their messages during the COVID-19 pandemic. Eight G7 leaders with active Twitter accounts were analyzed, namely: Joe Biden (USA), Justin Trudeau (Canada), Boris Johnson (UK), Yoshihide Suga (Japan), Emmanuel Macron (France), Charles Michel (EU Council), Ursula von der Leyen (EU Commission), and the Italian Presidency Council. A total of 9,918 COVID-19-related tweets were fetched over 16 months from April 2020 to July 2021. After the tweets were preprocessed, a three-stage analysis pipeline was applied: (1) preliminary descriptive analysis, (2) sentiment analysis based on VADER sentiment, Plutchik's emotions, and Profile of Mood States (POMS) emotions, and (3) topic discovery using Latent Dirichlet Allocation (LDA). The results showed that the G7 leaders were generally positive (69% positive, 12% neutral, 19% negative) regarding COVID-19. Trend analysis revealed that leaders' sentiments gradually became more positive as the pandemic progressed. While Justin Trudeau and Boris Johnson tweeted most frequently, Ursula von der Leyen was the most positive regarding COVID-19 topics. Topic modeling uncovered 9 main topics discussed by the leaders regarding COVID-19, namely: vaccination, situation monitoring and progress, negotiations, pandemic recovery, rebuilding future, community coordination, crisis support and subsidies, diplomacy, and global news. Among the topics, vaccination, diplomacy, and rebuilding future were the most positively discussed and vaccination was most closely associated with the emotion trust. The proposed topic classification of leaders' tweets was 77% accurate based on an ensemble predictive model. Drawing from the sentiments, topics, and frames that emerged from the Twitter usage of G7 leaders during COVID-19, world leaders and organizations can better frame crisis messaging and implement communication strategies to share information during pandemics.

Keywords: COVID-19, Sentiment Analysis, Natural Language Processing, Topic Modeling, Twitter, G7 World Leaders, Pandemic, Public Health

1. INTRODUCTION

The end of 2019 marked the onset of the highly contagious Coronavirus Disease 19 (COVID-19) which continues to pose a health threat to humanity. After the first cluster of the novel coronavirus was reported in Wuhan, China on 31 December 2019, the virus rapidly spread around the world, killing 4.4% of infected individuals within a few months (Lee, 2020). With the growing number of cases and deaths, COVID-19 quickly received global attention and was declared a pandemic by the World Health Organization (WHO) on 11 March 2020 (Haman, 2020). As a result, health organizations, government bodies, and world leaders increasingly turned to online platforms and social networking websites to inform the public on COVID-19 related matters such as safety measures, health protocols, and the governments' pandemic response. The leaders' choice of using online communication channels is rooted in the premise that people tend to spend more time on social media in crisis situations (Lamsal, 2021). As such, leaders can quickly communicate COVID-19 related information to a large proportion of the public by leveraging social media platforms as communication channels in addition to traditional media such as television, radio, and newspapers. With over 206 monetizable daily active users as of Q2 2021 (Statista, 2021), the microblogging social networking website Twitter emerged as a powerful public health communication tool for leaders to disseminate COVID-19 related information. Among the different social media platforms, Twitter is preferred by political leaders because people perceive and use the platform as a convenient medium to seek information, guidance, news, and social utility (Parmelee & Bichard, 2012).

Exploring the Twitter usage of all political leaders around the world may be technically feasible in terms of data collection; however, it would be an unnecessarily arduous and inefficient undertaking in terms of comparative analysis. Accordingly, this study specifically focuses on the Twitter usage of the Group of Seven (G7) world leaders in response to COVID-19. The Group of Seven (G7) is an informal group comprising the seven most highly developed democracies in the world, namely: Germany, Italy, France, Japan, the United Kingdom (UK), United States (USA), and Canada, as well as the European Union (EU). Leaders of the G7 member countries meet on an annual basis to discuss key global issues related to economic governance, energy policy, and international security (Joseph & Garside, 2010). While relatively small in number, the G7 countries collectively account for 60% of the global net wealth, 40% of the global GDP, and 10% of the global population (Credit Suisse, 2018). There are several advantages to the small size of the G7 relative to larger forums such as the G20 or United Nations (UN). For instance, leaders of the G7 countries can more easily achieve concrete results by engaging in informal and frank dialogue that draws from their shared values of democracy, human rights advocacy, and free markets. Moreover, the G7 represents a space where new global solutions and initiatives are piloted before being promoted to the G20, UN, and other global forums (France Diplomacy, 2018). Despite having no legal existence, the G7 possesses a collective, synergistic decision-making power that is impactful on a global scale. In this regard, the G7 is a small, but highly valuable subset of world leaders globally, making it a suitable representation for analysis.

Among the G7 leaders, only Angela Merkel, Chancellor of Germany, does not have an active Twitter account as of this writing and is therefore not included in the analysis. Table 1 shows the list of G7 leaders analyzed in this study. The eight selected G7 world leaders have a vast public reach on Twitter as depicted by their combined following of 64.8 million users.

Table 1. List of G7 Leaders Analyzed.

G7 Member/Leader	Designation	Verified Twitter Handle	Tweets[a]	Followers[b]
Joe Biden	President of the United States	@POTUS, @JoeBiden	2673	45.1M
Boris Johnson	Prime Minister of the United Kingdom	@BorisJohnson	983	3.6M
Justin Trudeau	Prime Minister of Canada	@JustinTrudeau	2060	5.7M
Emmanuel Macron	President of France	@EmmanuelMacron	917	7.2M
Yoshihide Suga	Prime Minister of Japan	@sugawitter	448	457k
Charles Michel	President of the EU Council	@eucopresident	574	1.2M
Ursula von der Leyen	President of the EU Commission	@vonderleyen	1432	639k
Palazzo Chigi	Presidency of the Council of Ministers of Italy	@Palazzo_Chigi	831	814k
			9918	**64.8M**

[a]*Number of COVID-19 related tweets from April 2020 through July 2021 based on retrieval criteria specified in Appendix A,* [b]*As of August 2021*

The language and framing used by G7 leaders in their online posts can influence the opinions, behaviors, and mental states of Twitter users all over the world, not just citizens of the G7 countries (Bernhardt, 2004). Given this influence, the leaders have a political, professional, and ethical responsibility to appropriately communicate correct information regarding COVID-19. This provides a strong rationale to investigate the Twitter usage and activity of the G7 leaders during COVID-19. Prior research in this area has been saturated towards the public sentiment of Twitter users during COVID-19; only a handful of studies have explored the Twitter usage of groups of individuals such as world or state leaders. Haman (2020) investigated the Twitter usage of UN countries' leaders during COVID-19 using regression analysis. Similarly, Rufai and Bunce (2020) explored the Twitter usage of G7 leaders in response to COVID-19, but their scope was limited to qualitative content analyses. This study employs sentiment analysis and topic modeling using Natural Language Processing (NLP) to explore how the G7 leaders used Twitter to communicate their messages during COVID-19. A deep, data-driven understanding of the tweet content of G7 leaders regarding COVID-19 would not only illuminate how political leaders use platforms such as Twitter to convey information during crises but also help inform public policy regarding crisis messaging and public health communication during pandemics. Specifically, the following research questions are addressed by this study:

1. How frequently did each G7 leader tweet about COVID-19 from April 2020 to July 2021?
2. What was the overall tweet sentiment of each G7 leader regarding COVID-19 and the topics surrounding it?
3. Which common words, collocations, and latent topics emerged from the tweet content of G7 leaders during COVID-19?
4. To what extent was the tweet sentiment of G7 leaders correlated with the development of COVID-19 in their respective countries?
5. How did G7 leaders' sentiments and tweet topics regarding COVID-19 change over time from April 2020 to July 2021?
6. With what predictive accuracy can G7 leaders' COVID-19 related tweets be classified into the identified topics?
    - 6.1 How do different machine learning algorithms compare in classifying the G7 leaders' tweets into topics and which algorithm performs best?
    - 6.2 How do the identified topics compare in terms of ease of classification?
    - 6.3 In terms of feature importance, which words contribute most towards differentiating the classified topics?

The structure of this paper is organized as follows. Section 1 has introduced the background, rationale, and research goals underpinning the study. Next, Section 2 presents a review of related literature and previous work. Section 3 then details the research methods, tools, and techniques used for data analysis. Section 4 summarizes the results, findings, and insights of the data analysis. Based on the findings in Section 4, Section 5 culminates the study with a brief discussion of policy implications, strengths and contributions, limitations, and recommendations for future work.

2. LITERATURE REVIEW

Use of Twitter Data for Analysis of Pandemics. A growing body of literature has been centered on measuring the efficacy of social media platforms for public health communication. Social media platforms such as Facebook and Twitter have proven to be effective communication mediums for health organizations and government leaders to disseminate COVID-19 related information. To this end, Twitter has an edge over other social media platforms in that it offers short, real-time content as well as access to networks of related conversations via hashtags (Su et al., 2021). As a microblogging platform, Twitter is ideal for research that requires opinion analysis as it enables users to freely disclose their location, thoughts, opinions, and feelings (Villavicencio et al., 2021). Prior to the COVID-19 pandemic, Twitter data has been successfully used to investigate dynamics of epidemics. For instance, Aramaki, Maskawa, and Morita (2011) used NLP to develop a sentiment-based method for detecting Influenza epidemics based on Twitter data. Twitter data was also extensively leveraged during public health outbreaks including Ebola, Zika virus, H1N1, and SARS for the prediction of disease spread, dissemination of public health information, and assessment of public views towards the outbreaks (Rufai & Bunce, 2020). Alamoodi et al., (2021) conducted a review of sentiment analysis studies regarding epidemics, pandemics, and disease outbreaks over a 10-year period and noted the importance of information dissemination during pandemics for improved response times and advanced planning. In an arguably short duration, the COVID-19 pandemic has been extensively studied using Twitter data. For instance, Twitter data has been used for the evaluation of restrictions such as social distancing, detection of self-reported COVID-19 symptoms, analysis of hashtags to discover psychological effects of lockdowns, and sentiment analysis of tweets (Su et al., 2021).

Twitter Sentiment Analysis Regarding COVID-19. Studies have applied Twitter sentiment analysis to explore different aspects and themes relating to COVID-19. Alhajji, Khalifah, and Aljubran (2017) analyzed tweets related to COVID-19 preventive measures in Saudi Arabia and found that tweets about religious restrictions were less negatively discussed whereas tweets about university closures and lockdowns had a higher negative sentiment. They further argued that sentiments regarding preventive measures tend to be influenced by the level of trust residents have towards their local government. A research by Barkur, Vibha, and Kamath (2020) scrutinized sentiments towards the nationwide lockdown in India. Their sentiment analysis found that Indians generally had positive sentiment towards the lockdown and displayed trust towards the government. The authors identified individual safety and the motivation to flatten the spread curve as main drivers of the prominent positive sentiment among Indians. Prominent negative sentiments that surfaced included fear, sadness, and disgust, although these emotions did not stand out as much as trust and positivity. Similar findings were obtained by Bhat et al., (2020) who argued that the prominence of positive sentiments emerged because people trusted and appreciated the efforts of government and frontline workers despite being stressed under lockdown. In addition, their analysis observed that 52% of tweets expressed positive sentiments, 34% neutral, and 14% negative. Twitter sentiment analysis studies on COVID-19 vary in regards to polarity results depending on the time-period covered, location, keywords, hashtags, sample size, and other factors. To illustrate, Boon-Itt and Skunkan (2020) obtained considerably different results compared to Bhat et al. (2020) and found that only 22% of tweets in their model had positive sentiments while 78% contained negative sentiments. Sengupta et al. (2021) investigated the overall impact of COVID-19-induced restrictions such as social distancing and lockdowns on mental health. They performed topic modeling using Latent Dirichlet Allocation (LDA) to better understand emotions and topics emerging from mental health-related tweets shared between June and July 2020. In their novel exploratory work, Sanders et al. (2020) used clustering and sentiment analysis techniques to unravel insights about mask-wearing attitudes and behaviors from a database of over one million tweets from March to July 2020. They found that sentimentality related to mask-wearing around the world became increasingly negative as the pandemic progressed. Boon-Itt and Skunkan (2020) looked into mentions and sentiments related to COVID-19 symptoms such as fever, cough, and fatigue and observed statistically similar distributions of symptom mentions in the Twitter data compared to actual medically recorded symptoms. Additionally, a handful of studies have proposed machine learning and deep learning models for the classification of COVID-19 related tweets. For example, Behl et al. (2021) built a Multilayer Perceptron (MLP) classifier to identify tweets about resource needs to match them with resource availability during the pandemic for the benefit of disaster relief operations.

Vaccination and Reopening. As the COVID-19 pandemic matures, studies are increasingly studying sentiments regarding vaccination and reopening as the volume of tweets surrounding these themes continues to grow proportionally. Villavicencio et al., (2021) performed a sentiment analysis specifically on tweets related to the rollout of COVID-19 vaccines in the Philippines. They mined tweets of both English and Filipino languages and found that 83% were favorable towards vaccination, 9% were indifferent, and 8% held negative sentiments. Analyzing sentiments related to vaccination can help illuminate the level of vaccine hesitancy in different countries and invite corrective action (Puri et al., 2021). Samuel et al., (2020) built a public sentiment analysis model associating emotions such as intense fear, confusion, trust, eagerness, and volatile emotions to various new normal scenarios as part of the movement to reopen the economy. Their findings suggest that tweets from American Twitter users reflected stronger positive sentiment support than negative sentiment support for reopening the US economy based on textual data analytics and statistical validation. On a more global scale, a study by Ahmed, Rabin, and Chowdhury (2020) observed low levels of negative sentiment towards reopening and found that the dominant emotion of fear was significantly weaker during reopening phases compared to the initial lockdowns in March 2020. Tourism and cross-border travel are also emerging as topics of research interest as the COVID-19 pandemic progresses. Lu and Zheng (2021) examined time-series dynamics of Twitter public sentiment regarding cruise tourism and the factors that influence it. They concluded that people may be more eager to travel and explore compared to pre-pandemic levels to as a result of extensive quarantine, isolation, and travel restrictions. As the pandemic matures, studies will increasingly apply Twitter sentiment analysis on COVID-19 tweets to study reopening, travel, and post-pandemic sentiments. Table 2 presents a curated list of related studies that have so far used Twitter Sentiment Analysis to study COVID-19 dynamics.

Table 2. Related Work on Twitter Sentiment Analysis during COVID-19

Study	Duration covered	Location	Tweet Authors	Tweet Languages	Keywords related to:	Primary Analytic Focus
Abd-Alrazaq et al. (2020)	02 Feb 2020 - 15 Mar 2020	Global	Public users	English	COVID-19	Sentiment analysis, topic modeling (LDA)
Ahmed et al. (2020)	03 May 2020 - 15 May 2020	USA	Public users	English	Reopening	Sentiment analysis, topic modeling (LDA)
Alhajji et al. (2020)	27 Feb 2020 - 03 Mar 2020	Saudi Arabia	Public users	Arabic	COVID-19 prevention	Sentiment analysis
Barkur, Vibha, & Kamath (2020)	25 Mar 2020 - 28 Mar 2020	India	Public users	English	COVID-19	Sentiment analysis
Boon-Itt & Skunkan (2020)	13 Dec 2019 - 09 Mar 2020	Global	Public users	English	COVID-19	Sentiment analysis, topic modeling (LDA)
Garcia & Berton (2021)	17 Apr 2020 - 08 Aug 2020	Global	Public users	English, Portuguese	COVID-19	Sentiment analysis, topic modeling (GSDMM)
Haman (2020)	30 Dec 2019 - 07 May 2020	U.N Member Countries	State Leaders	Multilingual	COVID-19	Linear regression
Jang et al. (2021)	21 Jan 2020 - 31 May 2020	Canada, USA	Public users	English	COVID-19	Aspect-based Sentiment analysis, topic modeling (LDA)
Kausar et al. (2021)	21 Jun 2020 - 20 Jul 2020	Top 11 Infected Countries	Public users	English	COVID-19	Sentiment analysis
Kruspe et al. (2020)	01 Dec 2019 - 31 Apr 2020	Europe	Public users	Multilingual	COVID-19	Sentiment analysis
Kumar, Khan, & Kalra (2020)	17 Mar 2020 - 30 Mar 2020	Global	Public users	English	COVID-19	Sentiment analysis
Lamsal (2021)	20 Mar 2020 - 17 Jul 2020	Global	Public users	English	COVID-19	Sentiment analysis, network analysis,
Lu & Zheng (2021)	01 Feb 2020 - 18 Jun 2020	Global	Public users	English	Cruise tourism	Sentiment analysis, topic modeling (LDA)
Manguri et al. (2020)	09 Apr 2020 - 15 Apr 2020	Global	Public users	English	COVID-19	Sentiment analysis
Müller et al. (2020)	12 Jan 2020 - 16 Apr 2020	Global	Public users	English	COVID-19	Sentiment analysis, CT-BERT novel NLP model for COVID-19
Pasteur & Koch (2020)	27 Feb 2020 - 25 Mar 2020	Global	Public users	English	COVID-19	Sentiment analysis, Association rules mining
Pokharel (2020)	21 May 2020 - 31 May 2020	Nepal	Public users	English	COVID-19	Sentiment analysis
Rahman et al. (2021)	30 Apr 2020 - 08 May 2020	USA	Public users	English	Reopening	Binary logit regression
Rufai & Bunce (2020)	17 Nov 2019 - 17 Mar 2020	G7 Countries	G7 Leaders	Multilingual	COVID-19	Qualitative content analysis
Samuel et al. (2020)	01 May 2020 - 09 May 2020	USA	Public users	English	Reopening	Sentiment analysis, scenario analysis
Sanders et al. (2020)	17 Mar 2020 - 27 Jul 2020	Global	Public users	English	COVID-19	Sentiment analysis, Text clustering (K-Means, t-SNE)
Sengupta et al. (2020)	28 June 2020 - 28 July 2020	Global	Public users	English	Mental health	Sentiment analysis, topic modeling (LDA)
Singh et al. (2020)	29 Mar 2020 - 31 Mar 2020	Global	Public users	English	COVID-19	Sentiment analysis, topic modeling (LDA)
Singh, Jakhar, & Pandey (2021)	20 Jan 2020 - 25 Apr 2020	Global vs India	Public users	English	COVID-19	Sentiment analysis (BERT)
Su et al. (2021)	23 Jan 2020 - 25 Mar 2020	USA	Public users	English	COVID-19	Topic modeling (LDA)
Villavicencio et al. (2021)	01 Mar 2021 - 31 Mar 2021	Philippines	Public users	English	COVID-19 Vaccines	Sentiment analysis
Xue et al. (2020)	07 Mar 2020 - 21 Apr 2020	Global	Public users	English	COVID-19	Sentiment analysis, topic modeling (LDA)
This study	01 Apr 2020 - 31 Jul 2021	G7 Countries	G7 Leaders	Multilingual	COVID-19	Sentiment analysis, topic modeling (LDA)

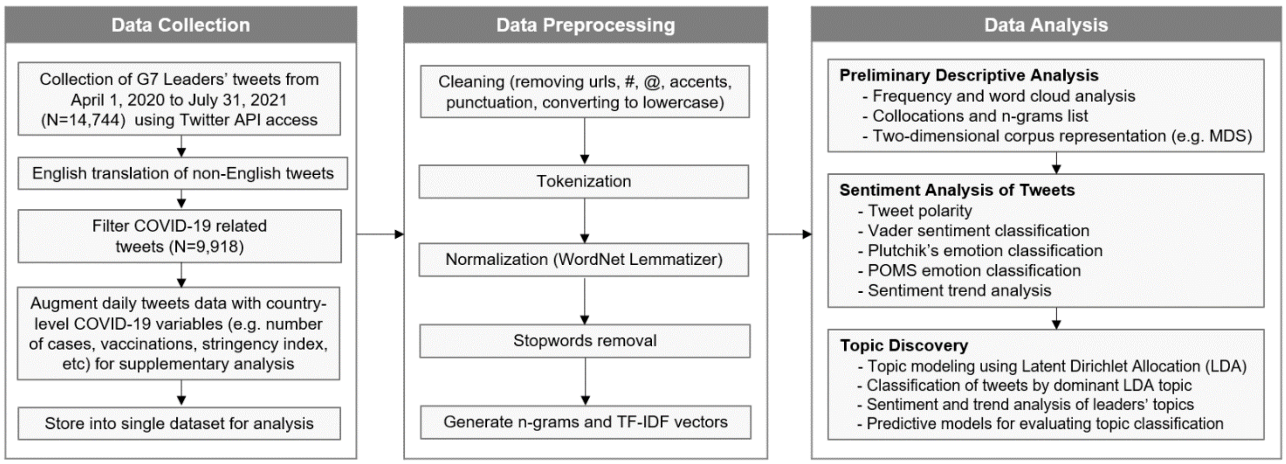
Twitter Usage of World Leaders During COVID-19. Given the high effectiveness of Twitter as an online tool for public health communication, governments and political leaders have recognized the platform as a powerful medium to share COVID-19 related information and engage with the public during the pandemic. In emergency situations such as pandemics, government leaders and organizations have an important responsibility to come forward and communicate accurate, complete, and timely information to shield against possible information credibility issues on microblogging platforms such as Twitter (Castillo, Mendoza, & Poblete, 2011). This imperative is accentuated by the fact that people tend to spend more time on social media in crisis situations, whether man-made or natural (Lamsal, 2021). While several studies have carried out Twitter sentiment analysis on COVID-19 related tweets to understand the sentiments of the general public, only two notable studies have looked into the Twitter usage of world leaders regarding COVID-19. In this regard, there exists an evident gap in the literature. Analyzing the tweets of representative world leader groups such as the G7, G20, or BRICS provides a good representation of the Twitter usage dynamics of world leaders around the world (Larionova & Kirton, 2020). First of its kind, a qualitative study by Rufai and Bunce (2020) featured a content analysis on the Twitter usage of G7 leaders in response to COVID-19. Their descriptive study compared tweet statistics among G7 leaders and classified the tweets as either informative, morale-boosting, or political. Haman (2020) conducted a similar study but considered tweets of 143 leaders of United Nations member states. Both these studies provided insightful descriptive and content analysis of the Twitter usage of world leaders, but neither performed sentiment analysis and unsupervised topic modeling on the tweet content using machine learning or NLP techniques. To fill this gap in the literature, this study builds upon previous literature and in part extends the work of Rufai and Bunce (2020) by applying sentiment analysis and topic modeling using NLP to analyze the Twitter usage of the G7 world leaders. In light of the related work summarized in Table 2, this study makes the following contributions to the literature:

- A large number of studies have shown interest in the global public sentiment of Twitter users during COVID-19. However, only two of the reviewed studies investigated the Twitter usage of world leaders during COVID-19, one of which performed a content analysis on G7 leaders' tweets (Haman, 2020; Rufai & Bunce, 2020). This is – to knowledge - the first work applying sentiment analysis and topic modeling using NLP on G7 world leaders' tweets during COVID-19. This study contributes to the literature on political leaders' Twitter usage for public health communication during pandemics like COVID-19.
- Studies in the literature have mostly considered under five months of Twitter data for their analyses. This research decision is perhaps driven by motivations to study a specific stage of the pandemic, data collection constraints due to large volume of data, or limited data availability for studies conducted early in the pandemic. This study considers a relatively longer time period of 16 months – allowing for trend analysis of the oscillation of sentiments and topics related to COVID-19 during this period.
- Most studies in the literature have limited their analysis to English language tweets and only a handful of studies have considered non-English tweets. For example, Garcia and Berton (2021) performed sentiment analysis and topic modeling on both English and Portuguese tweets. In this study, the multilingual tweets of G7 leaders are translated to English and included in the analysis to allow for a more precise and complete assessment.
- Finally, this study makes available the COVID-19 related Twitter data of G7 leaders from April 2020 to July 2021 which is collected, preprocessed, and collated into a single dataset in a structure that is conducive for analysis. The data can be used for further analysis or related research opportunities and will be shared upon request. In addition, the dataset structure and design may serve as a template to guide similar research involving Twitter sentiment analysis.

3. METHODS

Figure 1 shows a high-level process diagram of the research design used in this study. All data collection, processing, analysis, and visualization of results were performed using the software Python, Orange, SAS, and Microsoft Excel. Sample analysis workflows are shown in Appendix B.

Figure 1. Research Design



3.1. Data Collection

The Twitter API was used to fetch all tweets posted by the eight selected G7 leaders listed in Table 1 over a 16-month period from April 01, 2020 to July 31, 2021. A total of 14,744 tweets were retrieved from which close to two-thirds were in English (63.8%), followed by French (19.9%), Italian (8.6%), Japanese (4.4%), German (0.9%), and Spanish (0.4%). Table 3 shows the distribution of tweets by language.

Table 3. Distribution of Collected Tweets by Language

Tweet Language	Language Code	Tweet Count	Share of Total (%)
English	en	9406	63.8%
French	fr	2934	19.9%
Italian	it	1261	8.6%
Japanese	ja	655	4.4%
German	de	134	0.9%
Spanish	es	55	0.4%
Others	es	299	2.0%

All the 5,338 (36.2% of total) non-English tweets and retweets were translated into English using the Google Translate API. The French tweets mostly came from Emmanuel Macron and Justin Trudeau. Majority of the Italian tweets came from the account Palazzo_Chigi which represents the Presidency of the Council of Ministers of Italy. All the Japanese tweets came from Yoshihide Suga whose Twitter account and content is entirely in Japanese. German and Spanish tweets mostly came from Ursula von der Leyen. After translation, tweets about COVID-19 were identified based on a custom list of 99 keywords related to COVID-19 (See Appendix A). The list of keywords was largely sourced from the corpora of COVID-19 vocabulary compiled by Oxford English Dictionary (Wild, 2020). Tweets containing at least one instance of one or more words among the list of keywords were tagged as COVID-19-related and were retained for analysis. This filter resulted in a corpus of 9,918 tweets which meant 67% of the G7 leaders' tweets from April 2020 to July 2021 were related to COVID-19.

For supplementary analysis, the data collected from Twitter was then integrated with a number of additional country-specific COVID-19 variables from Our World in Data (OWID), made available by Mathieu et al. (2021). The motivation behind adding these variables was to examine whether the tweet activity and sentiments of the G7 leaders were, to any degree, influenced by the development of COVID-19 in their respective countries. The variables include daily-level data on the number of COVID-19 cases, vaccinations, reproduction rate, stringency index, COVID-induced anxiety, and hospital beds available for each G7 country. Several variables related to development such as GDP per capita, population density, GINI index, happiness index, corruption index, human development index, and social support index were also added to the dataset. For the European Union, average values for all the EU member countries were used. Appendix C contains the list of all variables in the final dataset used for analysis.

Figure 3. Integration of Twitter Data and OWID data

Twitter Data
date
country
month
year
day
month_year
weekday
username
leader
tweet
translated_tweet
language
length
like_count
retweet_count
author_status_count
author_favorites_count
author_following_count
author_follower_count

Country-level Indicators (OWID)
date
country
cty_new_cases_per_mil
cty_reproduction_rate
cty_new_vaccinations
cty_fully_vaccinated_per_hundred
cty_stringency_index
cty_covid_related_anxiety
cty_population_density
cty_gdp_per_capita
cty_hosp_beds_per_thousand
cty_human_development_index
cty_social_support_index
cty_corruption_index
cty_happiness_index
cty_GINI_index

3.2. Data Preprocessing

Unlike numerical data, textual data such as tweets is highly unstructured and must be preprocessed before analysis. Natural Language Processing (NLP) enables quantitative analysis of textual data. NLP provides a series of preprocessing techniques that can be applied to text to transform and prepare it for meaningful analysis. The most common NLP libraries and toolkits used in opinion mining and sentiment analysis literature include Natural Language Toolkit (NLTK) for Python, Fudan

NLP in Java, LTP in C++, Niu parser tool in C++, Gensim Python for topic modeling, Stanford CoreNLP in Java, SpaCy in Python, NLTK-based TextBlob, Tweepy, and OpenNLP (Solangi et al., 2018). For most studies in the literature, NLTK for Python is the preferred choice by researchers and scientists as it is flexible, backed by rich documentation, and encompasses a wide variety of NLP tasks required for text mining research (Bird, Klein, & Loper, 2009).

The preprocessing pipeline was programmatically implemented using Natural Language Toolkit (NLTK) and the text mining interface in Orange 3 which runs on common Python-based libraries including NLTK. First, the text was cleaned by converting all words to lowercase and removing all repeated words, punctuations, URLs, emails, and undesirable symbols and special characters (e.g. @, #, accents) from the tweets. Next, tokenization was performed using regular expression to break the text into smaller components (i.e. words). Third, normalization was applied using the WordNet Lemmatizer to reduce words to their base form; for example, 'traveling' and 'travels' were lemmatized to 'travel'. Normalization can be done either through stemming which essentially chops off word-endings or lemmatization which is often based on a dictionary for fuzzy word matching. Lemmatization was chosen as it is based on a dictionary and therefore more powerful (Bird, Klein, & Loper, 2009). The WordNet Lemmatizer, for instance, draws from a large lexical database to apply a network of cognitive synonyms to tokens (Fellbaum, 2005). Fourth, stopwords – which refer to high-frequency, non-meaningful words such as 'and', 'the', 'is' and irrelevant words such as 'RT', '&amp', etc. - were removed from the corpus. Next, n-grams were generated from the tokens. As a final step in data preparation, a document-term matrix was generated using a bag of words model to represent the corpus of tweets in a numerical form. The term frequency-inverse document frequency (TF-IDF), a statistical metric that quantifies the importance of each word in a corpus, was calculated for each document-term (i.e. tweet-word) combination to populate the matrix.

3.3. Data Analysis

Preliminary Descriptive Analysis. The corpus of tweets was first explored and visualized using descriptive statistics to gain an initial understanding of the data before carrying out sentiment analysis and topic discovery. Frequency analysis was performed; for instance, to understand tweet frequency per leader. Word cloud analysis was conducted to visually depict the most common words in the leaders' tweets. Collocation analysis was performed to draw out the most frequent n-grams such as unigrams, bigrams, and trigrams from the tweet content. Throughout the analysis, the data was visualized using a variety of multivariate statistical techniques for dimension reduction such as network analysis, multidimensional scaling (MDS), t-distributed stochastic neighbor embedding (t-SNE), isomap, uniform manifold approximation and projection (UMAP), principal component analysis (PCA), and correspondence analysis, to potentially draw out hidden patterns and insights from the tweets.

Sentiment Analysis. Sentiment analysis or opinion mining is a field of study that incorporates NLP to analyze sentiments, emotions, opinions, and attitudes in forms of written text such as product reviews, social media posts (e.g. tweets), blogs, or forum discussions (Liu, 2012). The primary goal of sentiment analysis is to understand whether the sentiment level in a piece of text is positive, neutral, or negative. The overall sentiment for each tweet was computed using the VADER lexicon and rule-based algorithm proposed by Hutto and Giblert (2014) which is optimized for conducting sentiment analysis on social media text. Using the VADER method, tweets were assigned a compound sentiment polarity score between -1 (most negative sentiment) to +1 (most positive sentiment.). Based on these polarity scores, a new column was appended to the data to label each tweet as either positive (polarity > 0), negative (polarity < 0), or neutral (polarity = 0). In addition to VADER sentiment, the emotion profile of the G7 leaders' tweets was determined using two popular classifications of emotions: Plutchik's wheel of emotions and Profile of Mood States (POMS). Proposed by Robert Plutchik in 1980, the Plutchik's classification of emotions is based on a wheel of eight emotions: joy, sadness, trust, disgust, fear, anger, surprise, and anticipation (Plutchik, 1980). Similarly, Profile of Mood States (POMS) is a psychological assessment of an individual's mood that uses six dimensions or categories to represent mood states: anger, fatigue, depression, confusion, tension, and vigor (Norcross et al., 1984). The Plutchik's and POMS emotion classifications were determined by subjecting the leaders' tweets to a pre-trained model developed by Colneric and Desmar (2020). To build their emotion recognition model, Colneric and Desmar (2020) first created a dataset of 73 billion tweets with emotion labels based on hashtags; for instance, a tweet containing the hashtag "#joy" was labeled with the emotion "joy". They then applied deep learning to train their emotion recognition models using 60% of the 73 billion tweets and achieved about 70% validation accuracy. Table 4 summarizes the three classifications of sentiment used in this study.

Table 4. Summary of Sentiment Classification Methods Used.

VADER Sentiment (Hutto & Gilbert, 2014)	Plutchik's Emotions (Colneric & Desmar, 2020)	Profile of Mood States (POMS) (Colneric & Desmar, 2020)
positive (polarity > 1)	Joy	Anger
neutral (polarity = 0)	Sadness	Depression
negative (polarity < 1)	Trust	Fatigue
	Disgust	Vigour
	Fear	Tension
	Anger	Confusion
	Surprise	
	Anticipation	

Among the three sentiment classification methods shown in Table 4, the VADER sentiment classification approach is the most commonly used in the literature and is the easiest to interpret as it simply assigns a sentiment polarity score to each tweet. From this polarity score, positive, negative, or neutral tweets can be determined based on predefined breakpoints (e.g. tweets with polarity > 0 may be assigned as "positive"). One advantage of the VADER method is that the breakpoints for classification can be set to different levels depending on the analytic goal. For instance, if the context of the text is naturally more positive, one may want to tweak the range of breakpoints and set a more stringent criterion (i.e. higher minimum polarity) to classify a tweet as positive. Furthermore, tweets are generally classified into three classes when using the VADER method; however, it is possible to create different levels based on the polarity. For example, it may be desirable in some situations to classify tweets as "very positive", "moderately positive", and so on, as opposed to just "positive". That said, the VADER method does not provide much insight regarding sentiment and emotion beyond the spectrum of positive to negative sentiment. Both Plutchik's and POMS classifications are based on pre-trained models and provide more depth into the emotion profile of the tweets through a larger number of classes. Plutchik's classification comprises eight emotions while POMS uses six mood-based emotions. However, one drawback of the two emotion-based methods is that the emotion classes tend to be imbalanced; for instance; Plutchik's tends to classify a disproportionately higher number of tweets as "joy" while the POMS classification gravitates towards "depression" (Colneric & Desmar, 2020). Consequently, this must be factored in during the interpretation of emotion classifications produced by the two methods. Most studies in the literature only use VADER or similar methods to classify tweets as positive or negative and only a few studies have explored the use of emotion-based classification methods. Among studies that have considered emotion classifications, Plutchik's method is more commonly used whereas the POMS classification has rarely been used. This study features both Plutchik's and POMS emotion classification methods which would help provide deeper insights regarding the emotion profile of the leaders' tweets.

Topic Discovery. Topic discovery using unsupervised machine learning was employed to formally detect and extract key topics in the G7 leaders' tweets. Specifically, the popular Latent Dirichlet Allocation (LDA) algorithm was used for topic modeling. During experimentation, other clustering methods such as K-Means clustering (See Appendix F), Latent Semantic Indexing, and HDBSCAN were also attempted to identify topics; however, LDA returned more distinct, meaningful, and interpretable topics. This coincides with the general consensus in the literature that LDA is currently the most favored technique by researchers for topic modeling on COVID-19 related tweets (Ahmed et al., 2020; Boon-Itt & Skunkan, 2020; Lu & Zheng, 2021; Jang et al, 2020). There are, however, a handful of studies that have applied other less-known methods. For example, Garcia & Berton (2020) used the Gibbs Sampling algorithm for the Dirichlet Multinomial Mixture (GSDMM) which is optimized for short pieces of text while Sanders et al., (2020) used a combination of K-Means and t-SNE to cluster topics. LDA is an unsupervised document classification approach that is able to find latent and natural topic groups within text. However, some characteristics of LDA may be disadvantageous such as having fixed and predetermined number of topics, uncorrelated topics, non-hierarchical structure, static and non-evolving topics, inability to allow supervised input, and indifference to sentence structure. Despite these downsides, LDA has been central to topic modeling and has advanced the field considerably (Liu et al., 2011). Rather than completely separating documents (i.e. tweets) into hard clusters, LDA treats each document as a probabilistic combination of topics and, likewise, each topic as a weighted combination of words (Boon-Itt & Skunkan, 2020). That is, a tweet can be associated with more than one topic and can overlap with other tweets in terms of topic content. In this study, LDA is applied as it performed best during experimentation. The LDA algorithm assigned per-topic probabilities to each tweet and the dominant topic was determined based on the most probable topic (i.e. topic with the highest loading). Furthermore, the optimal number (k) of LDA topics was determined based on the statistical measure - coherence score - which quantifies the similarity level of words within each topic. Coherence scores were computed for several topic models with sequentially different numbers of topics (k) and the most parsimonious topic model with the highest topic coherence beyond which the increase in coherence was negligible was chosen. Next, the LDA topics in the chosen model were manually interpreted and given meaningful names by inspecting the top keywords within each topic based on per-topic-per-word probabilities. The final set of topics were further analyzed in terms of their prevalence over time and their association with the G7 leaders and their sentiments. Finally, the predictive accuracy of the classified topics was validated using the TF-

IDF document-term matrix as features for the machine learning models. Seven machine learning algorithms were compared for this process, namely Logistic Regression, Stochastic Gradient Descent (SGD), Multilayer Perceptron (MLP), Support Vector Machine using Linear Kernel (Linear SVM), Naïve Bayes, k-Nearest-Neighbors (kNN), and a stacked ensemble model of the six individual algorithms. The best model was explained using SHAP (SHapley Additive exPlanations) values to identify the most important keywords in terms of their contribution towards topic classification. Developed by Lundberg and Lee (2017), SHAP is a method to explain feature importance based on Shapley values - a solution concept commonly used in cooperative game theory. SHAP was chosen over other methods of explaining feature importance such as LIME and DeepLIFT as it boasts higher computational performance and provides more consistent global interpretations (Lundberg & Lee, 2017).

4. DATA ANALYSIS AND RESULTS

4.1. Preliminary Descriptive Analysis

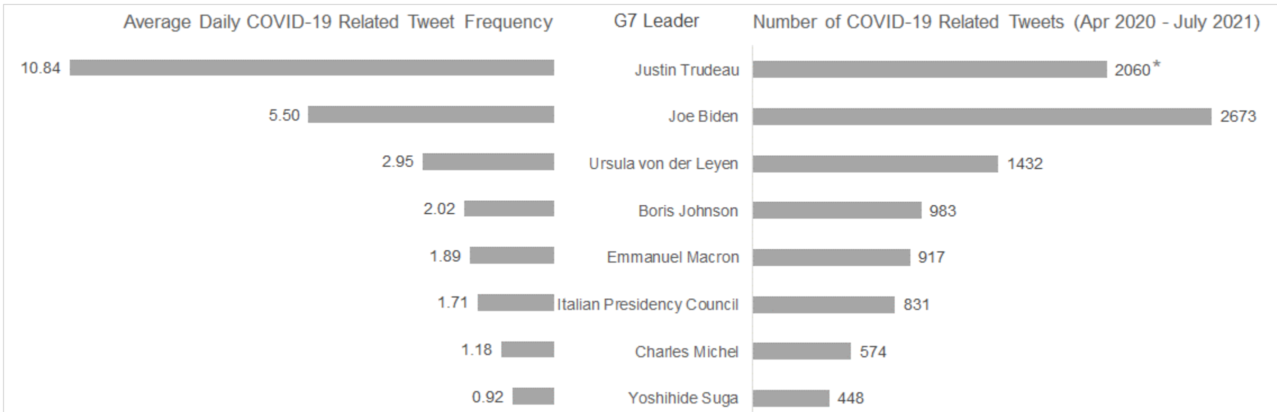
Table 5. Descriptive Statistics of Tweets by G7 Leader (Apr'20 – Jul'21)

G7 Leader	Tweets, sum	Tweet Length (translated), mean	Tweet Likes, mean	Retweets, mean
Boris Johnson	983	213	6006	890
Charles Michel	574	256	374	111
Emmanuel Macron	917	190	6561	1031
Italian Presidency Council	831	170	138	195
Joe Biden	2673	205	67201	10162
Justin Trudeau	2060*	218	1014	217
Ursula von der Leyen	1432	258	1075	271
Yoshihide Suga	448	338	1970	571
Overall	9918	221	19801	3055

* No tweets prior to Jan 2021 pulled for Justin Trudeau

This subsection presents a preliminary descriptive analysis of the leaders' tweets to build a foundational understanding of the data prior to sentiment analysis and topic modeling. Table 5 summarizes key descriptive statistics of the gathered data. A total of 9,918 COVID-19 related tweets were posted by the G7 world leaders from April 2020 to July 2021. All leaders were active over the covered duration except Justin Trudeau who had no tweets before January 22, 2021. Among the leaders, Joe Biden, President of the United States, had the highest number of tweets as he tweeted 2,673 times over the 16-month period. These tweets came from both his presidential (@POTUS) and personal (@JoeBiden) accounts. On the other hand, Yoshihide Suga, Prime Minister of Japan, was notably the least active, having tweeted only 448 times. In terms of tweet length, the Italian Presidency Council and Emmanuel Macron were the most succinct in their COVID-19 related posts as evidenced by their mean tweet length of 170 and 190 characters, respectively. While Yoshihide Suga tweeted least frequently, his tweets, after translation from Japanese, were the longest on average. Overall, the average tweet length of the G7 leaders was 221 characters. In terms of virality, Joe Biden's tweets were the most liked and retweeted whereas the tweets of Charles Michel and the Italian Presidency Council were the least liked and retweeted during the pandemic. This is largely influenced by each leader's number of followers on Twitter (See Table 1). For instance, Joe Biden has the biggest following so it is expected that his tweets would have the highest number of likes and retweets, on average.

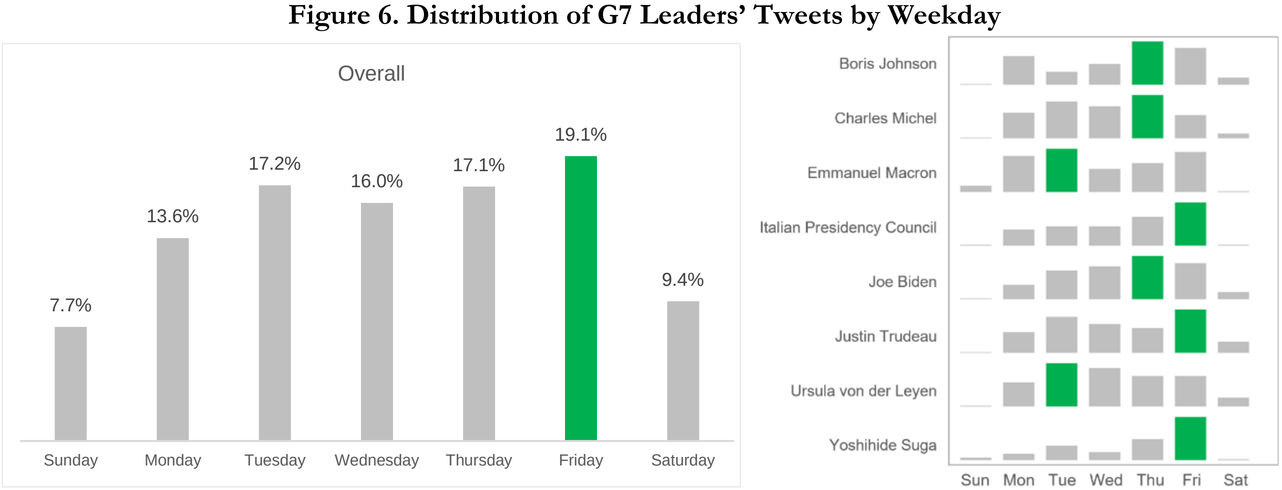
Figure 4. Tweet Frequency by G7 Leader (Apr 2020 – Jul 2021)



* No tweets prior to Jan 2021 for Justin Trudeau

Figure 4 shows the total and daily average tweet frequency of each G7 leader. While Joe Biden had the highest number of tweets, Justin Trudeau was the most active on Twitter as he tweeted 2060 times from January 2021 to July 2021, averaging 10.84 tweets per day. This is nearly double the frequency of Joe Biden who tweeted 5.5 times per day on average. The high variation in leaders' tweet frequency suggests that the country leaders have different communication styles towards using Twitter for posting COVID-19 related content. Justin Trudeau (Canada) and Joe Biden (USA) are evidently much more active on Twitter and leverage it to a higher degree to inform the public regarding COVID-19 compared to the other leaders.

Figure 6. Distribution of G7 Leaders' Tweets by Weekday



As depicted in Figure 6, the G7 leaders were generally more active on weekdays compared to Saturdays and Sundays. Overall, Friday had the highest share (19.1%) of COVID-19 related tweets. This pattern was consistent for all leaders. Relatively speaking, the chart also reveals that Yoshihide Suga had a much higher proportion of tweets on Fridays. A similar trend is observed for the Italian Presidency Council whose tweets appear to progressively increase over the week heading into Friday.

Figure 7. Word Cloud Representations of All G7 Leaders' Tweets

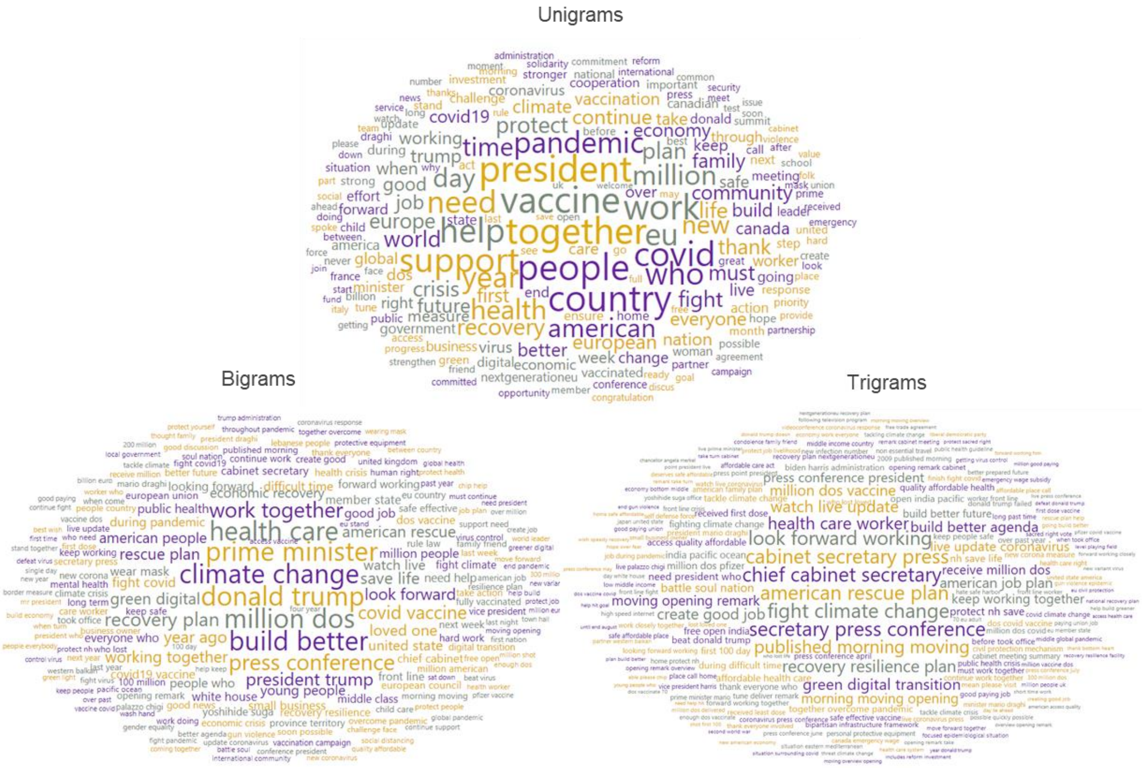


Figure 7 shows word cloud representations of the G7 leaders' tweet content. The top-most word cloud of unigrams reveals that "covid", "vaccine", "together", and "support" were some of the most commonly used words by the leaders. The word clouds of bigrams and trigrams provide more semantic context. For example, bigrams such as "covid vaccine" and "million dos" (i.e. normalized form of "million doses") are related to vaccination. Also apparent are bigrams such as "build better", "recovery plan", "fight covid", and "work together" and trigrams like "recovery resilience plan" which are related to pandemic response and recovery. Table 6 details the top 20 unigrams, bigrams, and trigrams based on their overall frequency in the corpus of tweets.

Table 6. Top 20 N-grams by Frequency

unigram		bigram		trigram	
together	1254	donald trump	286	american rescue plan	99
people	1231	climate change	233	chief cabinet secretary	66
vaccine	1205	million dos	202	fight climate change	56
president	1087	prime minister	197	cabinet secretary press	50
country	1072	build better	187	secretary press conference	50
work	1023	health care	184	look forward working	48
support	1003	press conference	150	published morning moving	44
covid	998	work together	143	health care worker	42
help	985	recovery plan	127	recovery resilience plan	41
need	907	covid vaccine	120	watch live update	40
who	895	president trump	115	green digital transition	38
eu	873	save life	106	create good job	37
year	853	year ago	104	million dos vaccine	36
pandemic	840	look forward	102	build better agenda	35
million	787	green digital	101	american job plan	34
health	767	rescue plan	100	morning moving opening	33
american	738	working together	100	moving opening remark	33
day	738	american rescue	99	press conference president	32
new	727	loved one	96	live update coronavirus	32
time	707	american people	89	receive million dos	31

Figure 8. Word Cloud Representation for each G7 Leader's Tweets

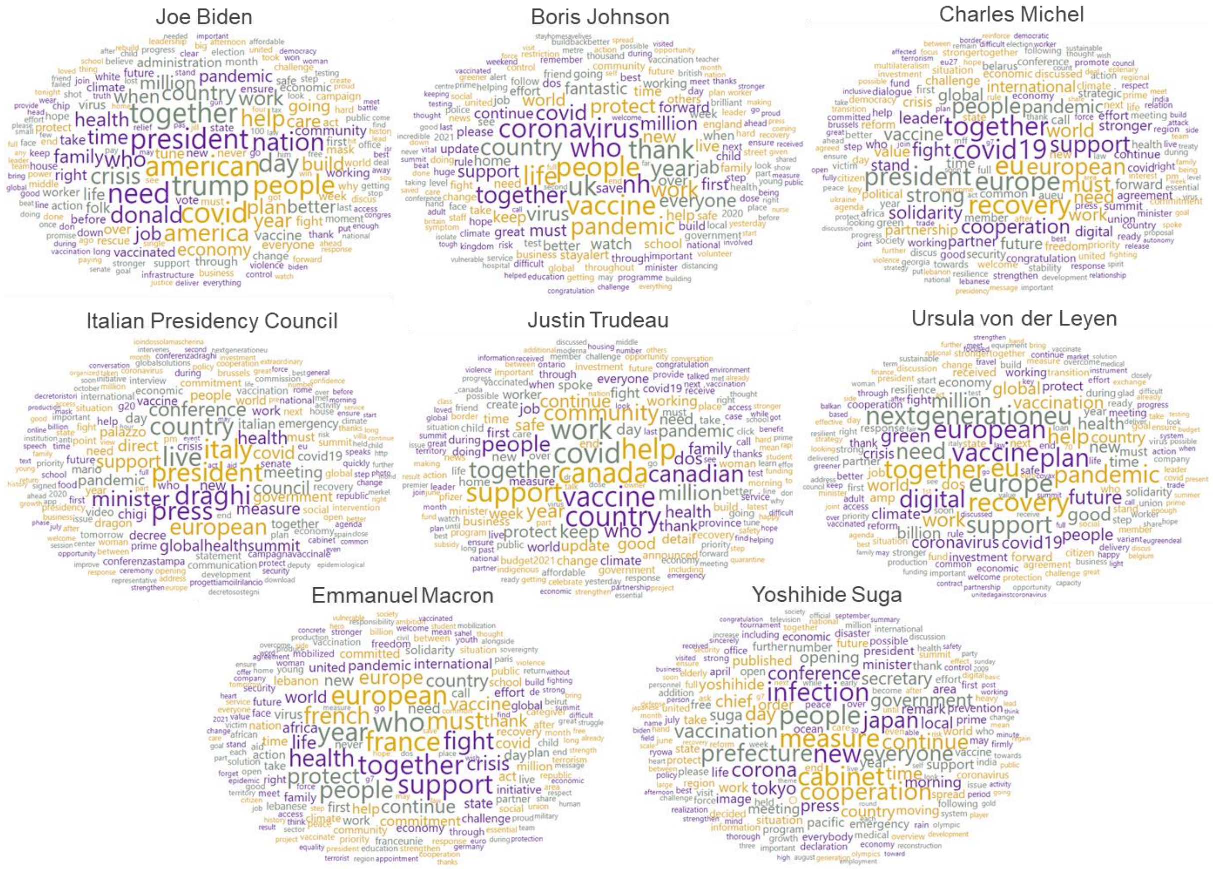


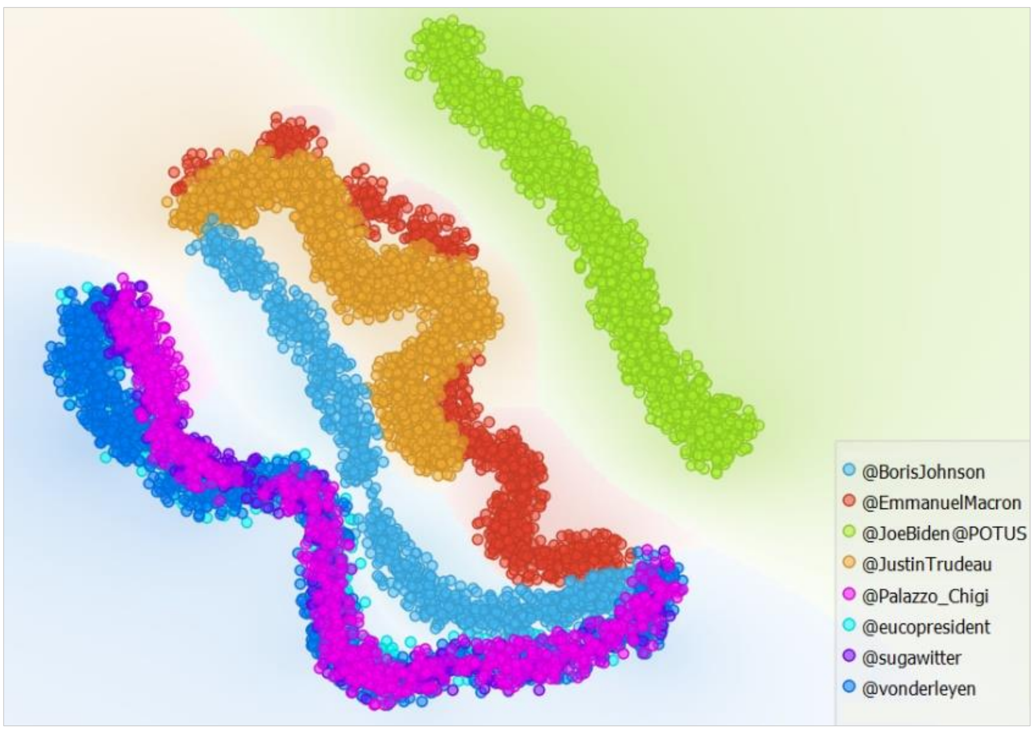
Figure 8 shows word cloud representations of each G7 leader's tweets. As hypothesized, a common theme among all leaders is the frequent mention of COVID-19 related words such as "vaccine", "support", "recovery", and "together". Several county-specific keywords are also apparent. For instance, the frequent mention of "prefecture" and "japan" by Yoshihide Suga, "france" and "french" by Emmanuel Macron, "trump" and "america" by Joe Biden, "eu", "europe", and "global" by both Charles Michel and Ursula von der Leyen, "canada" by Justin Trudeau, and "italy and "draghi" by the Italian Presidency Council. Table 7 lists the top 20 most common words by leader.

Table 7. Top 20 Unigrams by G7 Leader

Boris Johnson	Charles Michel	Emmanuel Macron	Italian Presidency Council	Joe Biden	Justin Trudeau	Ursula von der Leyen	Yoshihide Suga
people	eu	france	president	american	canada	eu	measure
who	covid19	together	draghi	president	vaccine	vaccine	new
uk	europe	who	live	trump	covid	europe	people
vaccine	recovery	european	italy	covid	help	recovery	cabinet
coronavirus	together	support	press	need	country	european	infection
thank	president	year	country	together	work	nextgenerationeu	prefecture
life	european	must	european	people	support	together	cooperation
nh	support	fight	minister	nation	canadian	plan	japan
country	must	people	conference	day	together	support	everyone
pandemic	people	french	support	america	community	pandemic	continue
together	strong	health	health	donald	people	digital	corona
year	cooperation	protect	council	crisis	million	need	vaccination
work	world	europe	covid	help	year	help	day
covid	fight	crisis	measure	who	who	future	conference
virus	need	continue	meeting	country	keep	million	government
protect	pandemic	vaccine	direct	plan	continue	work	time
million	vaccine	life	globalhealthsummit	time	dos	green	tokyo
everyone	solidarity	country	pandemic	work	pandemic	good	secretary
new	work	thank	government	year	safe	vaccination	chief
support	leader	world	new	health	health	covid19	press

The word clouds and list of top 20 unigrams per leader have made evident similarities and differences in keywords within the COVD-19 related tweets of the world leaders. Similar words are primarily related to COVID-19 issues whereas the differences stem from either country-specific words and topics or the word choice and usage of each leader. For example, the presence of "strong", "fight", and "must" among the top 20 words of Charles Michel perhaps suggests a more assertive style of written communication. Ursula von der Leyen frequently made use of more optimistic and future-oriented keywords such as "future", "recovery", "good", "green", and "digital". These observations can be statistically validated through polarity-based sentiment analysis in the next section. Moreover, the frequent use of words such as "need", "plan", "crisis", "together", and "help" by Joe Biden could suggest a communication or leadership style characterized by a greater sense of urgency and initiative regarding COVID-19 response. This may be linked to Biden's pandemic response plan which Maxmen and Subbaraman (2021) observed to be widely discussed on Twitter not only by Joe Biden but also his followers and researchers; some of whom regarded the plan as thorough, while others, ambitious.

Figure 9. Two-dimensional Projection of Tweets using t-SNE



To further explore the disparity in tweet content of the G7 leaders beyond subjective judgment, all the 9,918 tweets were projected into a two-dimensional space as shown in Figure 9. The motivation behind projecting the tweets in a low-dimensional space in this way was to see whether the previously identified country-specific and style-driven differences in the leaders' tweet content manifest themselves in a numerical or statistical sense. The two-dimensional t-SNE visualization of the G7 leaders' tweets was mapped using a bag of words model comprising of TF-IDF word vectors. Overall, the colored groups of leaders' tweets in the visualization indeed suggest statistical differences among G7 leaders based on word content. Joe Biden's tweets (green) are grouped relatively distant from the other leaders' tweets. There appears to be an overlap between the red and orange points which refer to Emmanuel

Macron and Justin Trudeau's tweets, respectively. Furthermore, the apparent overlap between the points of Charles Michel (light blue) and Ursula von der Leyen (dark blue) is expected given both represent the European Union and are therefore likely to coincide in terms of tweet content.

Figure 10. Word Co-occurrence Network Diagram of Top 50 Words in G7 Leaders' Tweets

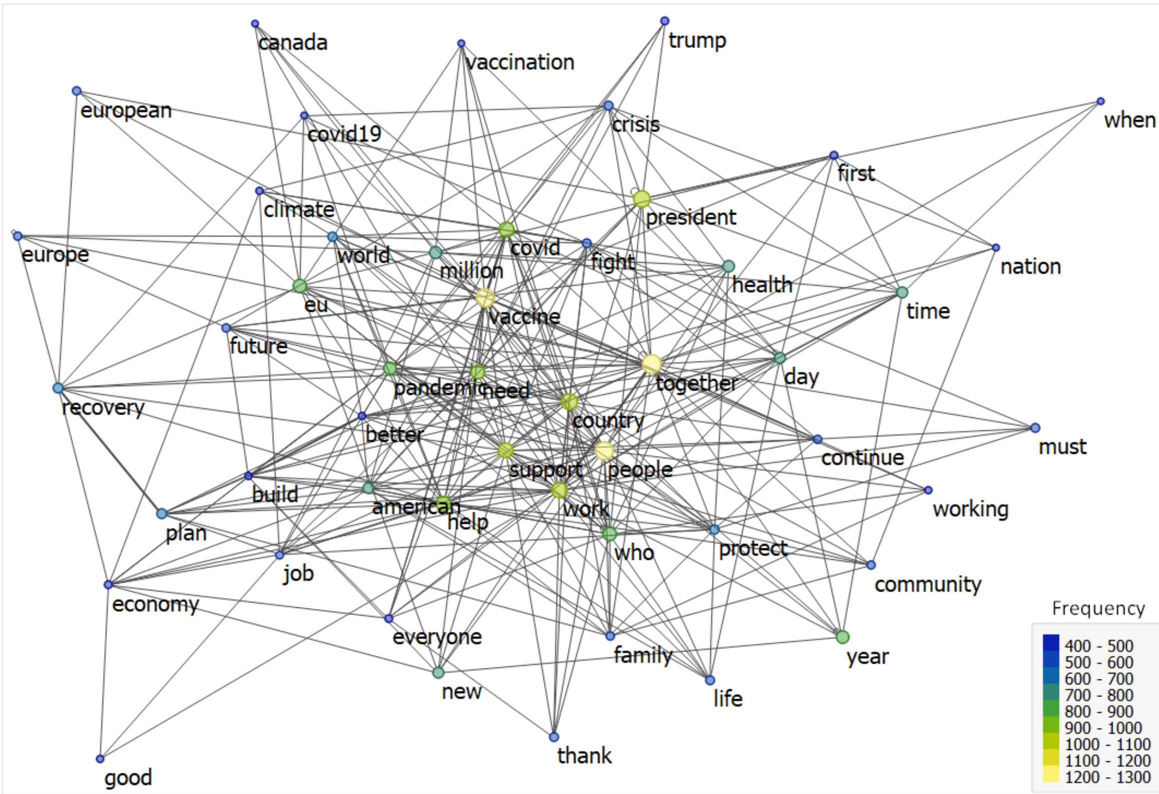
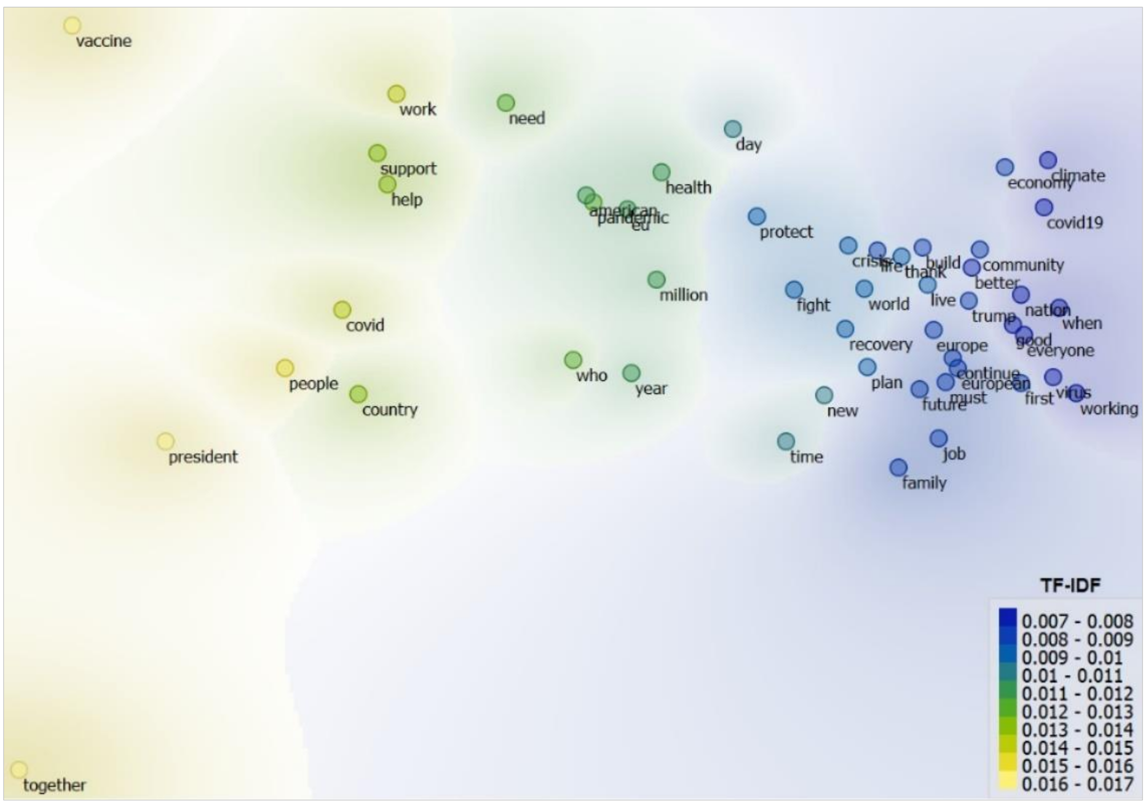


Figure 10 presents a network diagram to show potential interconnections between the top 50 most common words found in the G7 leaders' tweets. Theoretically speaking, if each of the 50 words/nodes was connected to (i.e. co-occurred with) every other word, there would be a maximum of 1,225 (i.e. (n(n-1)/2; 50(49)/2) edges linking the nodes to represent co-occurring n-grams. In the network diagram above, there are 312 edges or 6.24 co-occurring words per node which is much less than 1,225 edges, but the said theoretical possibility suggests how informative a network diagram can be. Moreover, the level of co-occurrence and similarity of the words can also be judged based on the length of the edges and the placement of the words in the two-dimensional space. In these ways, a network diagram can be much more powerful than a simplified inspection of a list of n-grams or word clouds. When two nodes are linked in the network diagram, it means that the two words represented by their respective nodes appear together within one or more tweets. If they co-occur with high frequency, the length of the edge linking them together would be shorter. That said, two co-occurring words may not always be bigrams, while a bigram is always a consecutive co-occurrence of two words. In this regard, a network diagram can unravel relationships between words that are not made clear by bigrams. Some observations from the network diagram include:

- Among the top 50 words, the word "good" co-occurs most with three other words, "economy", "job", and "time", as in "good economy", "good job", and "good time".

- "Recovery' co-occurs frequently with "plan" as does "build" with "future". Moreover, the close placement of the pairs of words "recovery plan", "build future" and "good economy" supports the intuitive association between the three relating to recovering and rebuilding from the pandemic. This observation presents a topic hypothesis that can be tested further with formal topic modeling using LDA.
- The word "community" co-occurs closely with "nation", "together", "country", "support", "protect", and "family". All these words represent a sense of community and cooperation and appear to be grouped close to each other. This again may be indicative of a theme or topic within the tweets.
- The most frequent words in the network such as "covid", "vaccine", "million", and "together" are pushed to the center and colored in yellow. Words in this group are linked closely to each other and represent the most important COVID-19 matters such as vaccination.

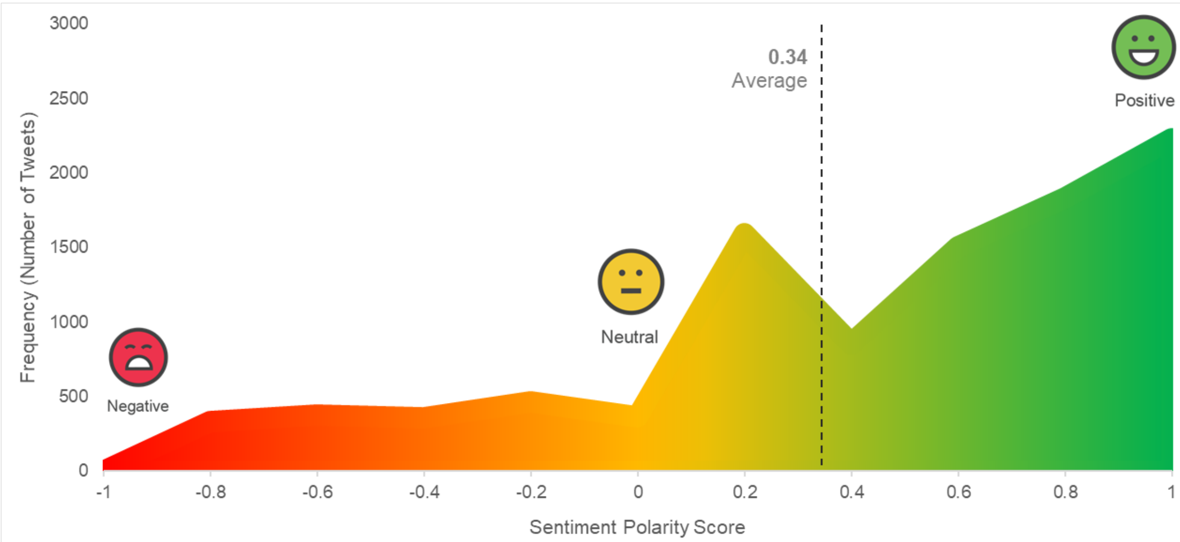
Figure 11. MDS Map of Top 50 Words in G7 Leaders' Tweets



The visualization above projects the top 50 words into two dimensions using multidimensional scaling (MDS). The words are arranged from highest to lowest TF-IDF frequency from left (yellow) to right (blue). Most frequent words such as "vaccine", "together", and "president" appear distant from the other words. Similar to the network diagram, related pairs of words such as "recovery" and "plan", "support" and "help", "climate" and "economy", "community" and "nation" are positioned in close proximity. Close groups of words are also apparent; for instance, "job", "family", and "future" may be related to fundamental priorities for survival during the pandemic.

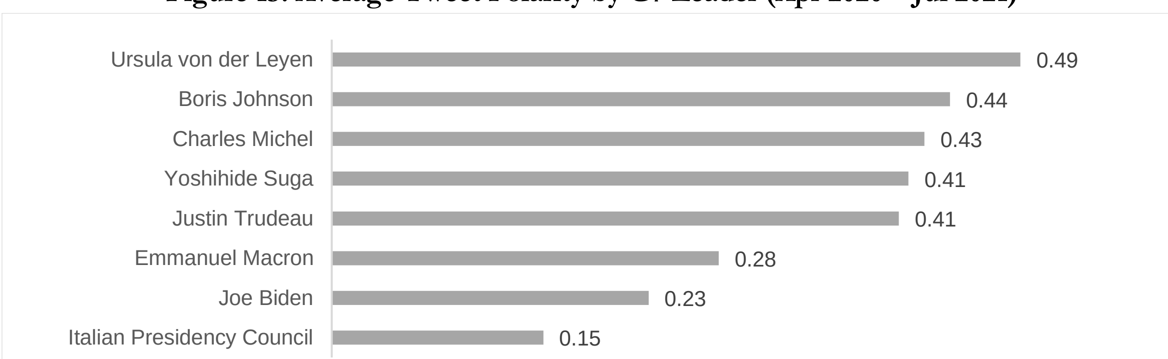
4.2. Sentiment Analysis

Figure 12. Polarity Histogram of G7 Leaders' Tweets



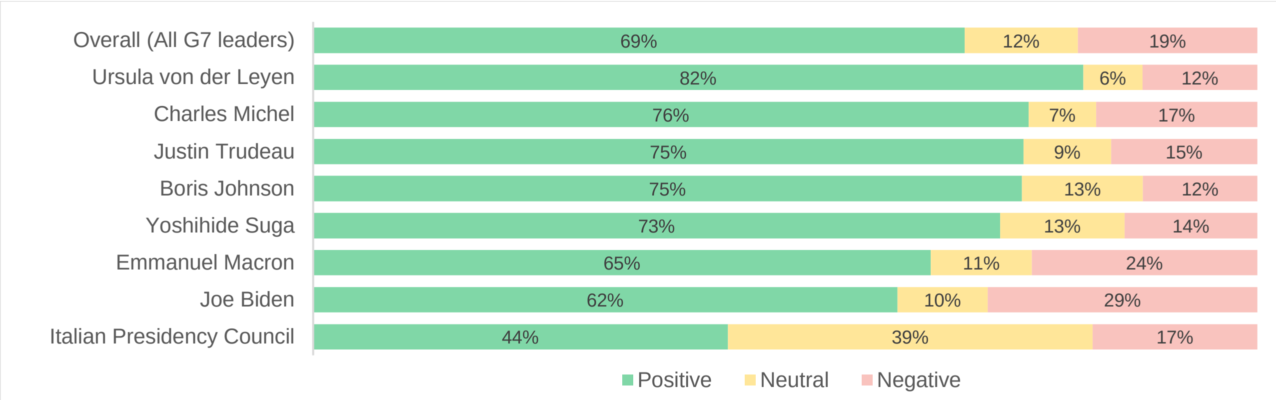
Building on the insights from the preliminary descriptive analysis, this subsection summarizes the results from the sentiment analyses applied to the G7 leaders' tweets regarding COVID-19. The average sentiment polarity across the G7 leaders' tweets was found to be 0.34 as shown in Figure 12. The distribution of the histogram reveals that the leaders' sentiments regarding COVID-19 matters were generally positive.

Figure 13. Average Tweet Polarity by G7 Leader (Apr 2020 – Jul 2021)



The average polarity of the leaders ranged from 0.15 to 0.49 as ranked in Figure 12. Among the G7 leaders, Ursula von der Leyen had the highest average polarity score (0.49) and was most positive regarding COVID-19 matters. This supports the observation from the network diagram where it was noted that she used more optimistic and future-oriented keywords compared to other leaders. On the other hand, the Italian Presidency Council had the lowest average polarity (0.15). Since a polarity of greater than 0 indicates positive sentiment, it can be inferred that none of the leaders had negative sentiments, on average, regarding COVID-19. The degree to which each leader was positive, negative, or neutral can be further explored by looking at the sentiment distribution.

Figure 14. Sentiment Distribution by G7 Leader



Overall, the sentiment distribution of all G7 leaders regarding COVID-19 tweets was 69% positive, 12% neutral, and 19% negative as shown in Figure 14. Ursula von der Leyen was the most positive in that 82% of her COVID-19 related tweets were positive while only 12% were negative. Along with Ursula von der Leyen, Boris Johnson was the least negative as he had the lowest proportion of negative tweets (12%). On the other hand, Joe Biden was the most negative regarding COVID-19 as he had the highest proportion of negative tweets (29%). Moreover, it can be seen that the Italian Presidency Council was the most neutral in sentiment compared to the other G7 leaders as 39% of their COVID-19-related tweets were neutral.

Figure 15. Monthly Sentiment Polarity Trend of G7 Leaders' Tweets

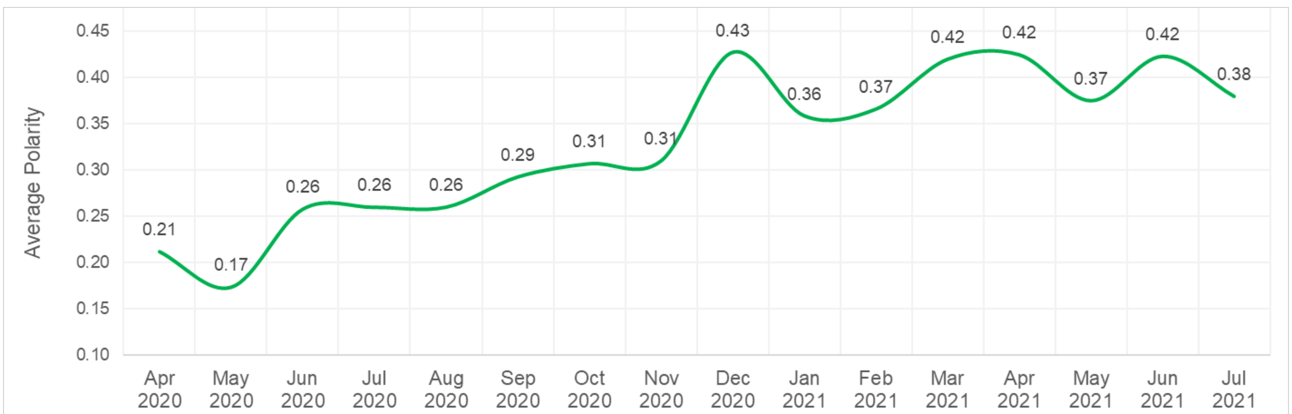
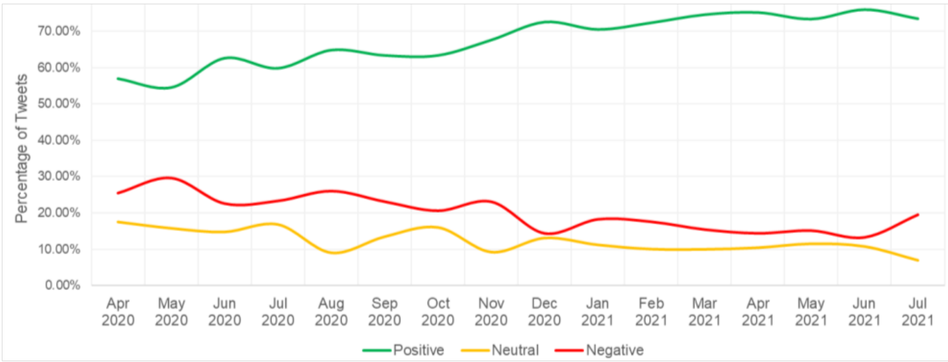


Figure 15 shows an upward trend in the sentiment of the leaders. The average sentiment polarity of the leaders' tweets was lowest in April and May 2020. During this period, the global number of infections reached 1,000,000, the global death toll surpassed 100,000, and more than a third of the world was under some sort of lockdown (Neilson & Woodward, 2020). The polarity steadily increased and became more positive for the remainder of the year as several countries began to ease lockdowns. Over the 16-month period covered, the sentiment polarity peaked in December 2020. This is perhaps underpinned by two reasons. First, the Pfizer and Moderna vaccines were authorized by the FDA and the UK in December 2020. In the context of the pandemic, this was a big positive respite amidst the growing death toll and fluctuating lockdowns at the time. Second, December represents a holiday season globally, possibly contributing to the higher (i.e. more positive) polarity in the tweet content of the leaders. Following the peak in December 2020, the polarity remained fairly constant throughout 2021, ranging from 0.36 to 0.43. Compared to 2020, the average polarity was much higher in 2021 compared to 2020, which may be driven by the proliferation of vaccination, declining number of cases, and easing of restrictions around the world during the first half of 2021 (Neilson & Woodward, 2020; Kantis, Kiernan, & Bardi, 2021).

Figure 16. Trend of Sentiment Distribution



As shown in Figure 16, the prevalence of leaders' positive tweets increased over time from April 2020 to July 2021 whereas the prevalence of neutral and negative tweets decreased. To explore the polarity trend at a deeper level of granularity, Figure 17 shows the daily average polarity of the leaders between April 2020 to July 2021.

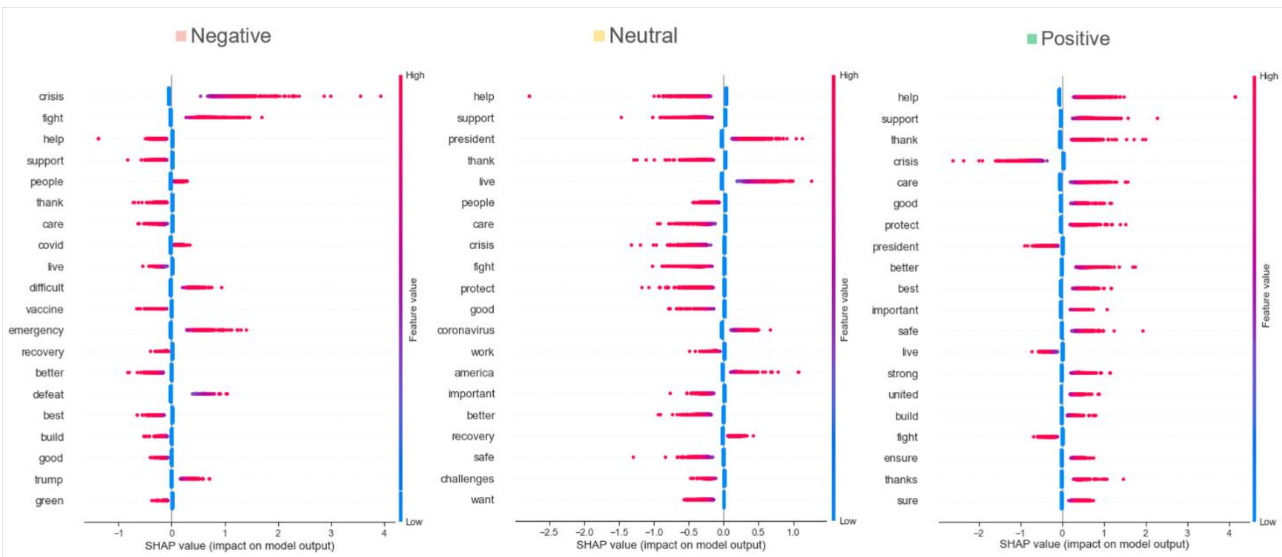
Figure 17. Daily Sentiment Polarity of G7 Leaders' Tweets (Apr 2020 – Jul 2021)

	Apr '20	May '20	Jun '20	Jul '20	Aug '20	Sep '20	Oct '20	Nov '20	Dec '20	Jan '21	Feb '21	Mar '21	Apr '21	May '21	Jun '21	Jul '21
1	-0.03	-0.02	0.05	0.52	0.22	0.22	0.37	0.14	0.45	0.54	0.27	0.48	0.50	0.42	0.20	0.45
2	0.19	0.04	0.29	0.24	0.19	0.28	0.26	0.16	0.44	-0.28	0.54	0.38	0.33	0.44	0.47	0.41
3	0.18	0.37	0.29	0.00	0.18	0.39	0.55	0.07	0.50	0.18	0.27	0.50	0.42	0.59	0.39	0.19
4	0.27	0.28	0.29	0.53	-0.05	0.34	0.54	0.01	0.43	0.53	0.24	0.21	0.82	0.30	0.38	0.20
5	0.37	0.38	0.22	0.13	0.09	0.25	0.36	0.09	0.20	0.25	0.39	0.48	0.46	0.52	0.50	0.34
6	0.31	0.08	0.31	0.27	0.41	0.09	0.40	0.51	0.38	0.40	0.60	0.47	0.39	0.45	0.48	0.28
7	0.51	0.11	0.10	0.43	0.06	0.24	0.45	0.51	0.48	0.20	0.36	0.39	0.50	0.46	0.20	0.26
8	0.25	0.41	-0.19	0.32	0.52	0.44	0.41	0.43	0.49	0.39	0.64	0.39	0.43	0.54	0.33	0.27
9	0.17	0.28	0.18	0.42	0.34	0.44	0.29	0.23	0.42	0.43	0.39	0.63	0.27	0.77	0.16	0.43
10	0.15	0.32	-0.05	0.18	-0.11	0.19	0.34	0.40	0.51	0.56	0.41	0.36	0.28	0.54	0.45	0.04
11	-0.08	0.37	0.27	0.14	0.44	0.21	0.19	0.25	0.48	0.41	0.46	0.40	0.68	0.35	0.50	0.45
12	0.12	0.26	-0.05	0.20	0.19	0.39	0.31	0.34	0.41	0.59	0.37	0.38	0.69	0.40	0.45	0.36
13	-0.08	0.19	0.12	-0.05	0.40	0.18	0.35	0.38	0.24	0.45	0.41	-0.10	0.46	0.28	0.58	0.38
14	0.29	-0.06	0.40	0.03	0.39	0.16	0.30	0.33	0.42	0.52	0.10	0.43	0.41	0.16	0.49	0.60
15	0.19	0.07	0.24	0.32	0.26	0.20	0.31	0.30	0.50	0.30	0.62	0.60	0.60	0.28	0.54	0.42
16	0.18	0.44	0.18	0.16	0.23	0.25	0.28	0.58	0.42	0.03	0.03	0.56	0.40	0.60	0.41	0.41
17	0.18	-0.05	0.19	0.15	0.25	0.33	0.42	0.87	0.53	0.69	0.09	0.61	0.53	0.29	0.31	0.09
18	0.09	0.09	0.45	0.30	0.27	0.21	0.14	0.56	0.37	0.33	0.54	0.42	0.46	0.33	0.37	0.36
19	0.46	0.22	0.25	0.12	0.36	0.76	0.15	0.31	0.23	0.17	0.43	0.34	0.26	0.37	0.37	0.45
20	0.16	0.17	0.13	-0.02	0.21	0.46	0.35	0.29	-0.16	0.38	0.28	0.45	0.50	0.40	0.19	0.43
21	0.18	0.23	0.26	0.32	0.35	0.62	0.25	0.50	0.47	0.31	0.39	0.49	0.49	0.28	0.48	0.43
22	0.16	0.13	0.39	0.35	0.42	0.06	0.35	0.36	0.22	0.46	0.06	0.21	0.30	0.56	0.65	0.27
23	0.18	-0.05	0.31	0.33	0.21	0.54	0.28	0.31	0.30	0.22	0.43	0.27	0.38	0.24	0.29	0.38
24	0.13	0.00	0.22	0.54	0.24	0.21	0.43	0.41	0.65	0.54	0.29	0.57	0.49	0.28	0.40	0.58
25	0.45	0.18	0.22	0.21	0.35	0.32	0.47	-0.14	0.78	0.46	0.44	0.52	0.30	0.30	0.44	0.61
26	-0.06	0.05	0.44	0.72	0.43	0.25	0.36	0.39	0.07	0.26	0.41	0.50	0.43	0.41	0.53	0.42
27	0.29	0.21	0.38	0.43	-0.03	0.38	0.31	0.37	0.22	-0.07	0.31	0.34	0.44	0.25	0.47	0.39
28	0.05	-0.16	0.12	0.30	0.23	0.18	0.26	0.68	0.66	0.43	0.15	0.13	0.49	0.31	0.49	0.44
29	0.30	0.12	0.33	0.22	0.19	0.38	0.25	0.54	-0.09	0.41		0.17	0.38	0.35	0.36	0.40
30	0.29	-0.01	0.31	0.45	0.31	0.25	0.33	0.36	0.52	0.43		0.49	0.37	-0.02	0.44	0.41
31		-0.27		0.31	0.22		0.01		0.66	0.55		0.52		0.51		0.33

-1 ---- polarity ---- 1

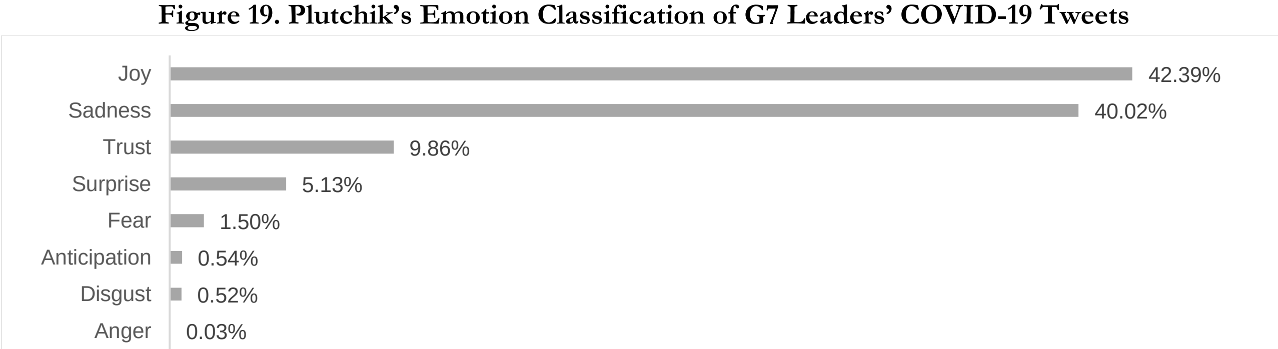
Days in green represent positive sentiment while days in red represent negative sentiment. The shift of the color scale from dominantly red on the left side of the heatmap to green towards the right side of the map is an indication that the leaders' sentiments became more positive in 2021 compared to 2020. Upon closer inspection, it can be seen that the average polarity on the 16th, 17th, and 18th of November was relatively more positive compared to the rest of the month. On these dates, interim results were released by both Moderna (on November 16) and Pfizer (on November 18) that their vaccine was 94% effective (Kantis, Kiernan, & Bardi, 2021). These developments may have contributed to the higher sentiment polarity of world leaders on these dates. Furthermore, most days in December, which represents festive season, appear to reflect high polarity levels. For example, Christmas day (December 25, 2020) had the most positive polarity (0.78) in December 2020. However, three days in December had negative polarity including December 20 which had the lowest polarity (-0.16) in the month. On December 20, several EU countries including France, Germany, Italy, Netherlands, and Ireland imposed travel restrictions on the UK (Kantis, Kiernan, & Bardi, 2021). While tweet sentiment can be affected by a myriad of factors, these findings suggest that leaders' tweet sentiments were possibly influenced by external COVID-19 related developments. To determine what influenced the positive, negative, and neutral tweets of the G7 leaders, it would be meaningful to examine keywords or unigrams that contributed most towards the classification of tweets in terms of feature importance. This is shown in Figure 18 below.

Figure 18. Top 20 Most Important Words per Vader Sentiment Class



Based on SHAP values for feature importance, the words "help", "support", "thank", "care", "good", "protect", "better", "safe", "strong", and "united" were the most positive words in the sense that they had the highest contribution towards classifying tweets as positive. Conversely, words such as "crisis", "emergency", "trump", "defeat", "fight", "difficult", and "defeat" were the most negative.

Figure 19. Plutchik's Emotion Classification of G7 Leaders' COVID-19 Tweets



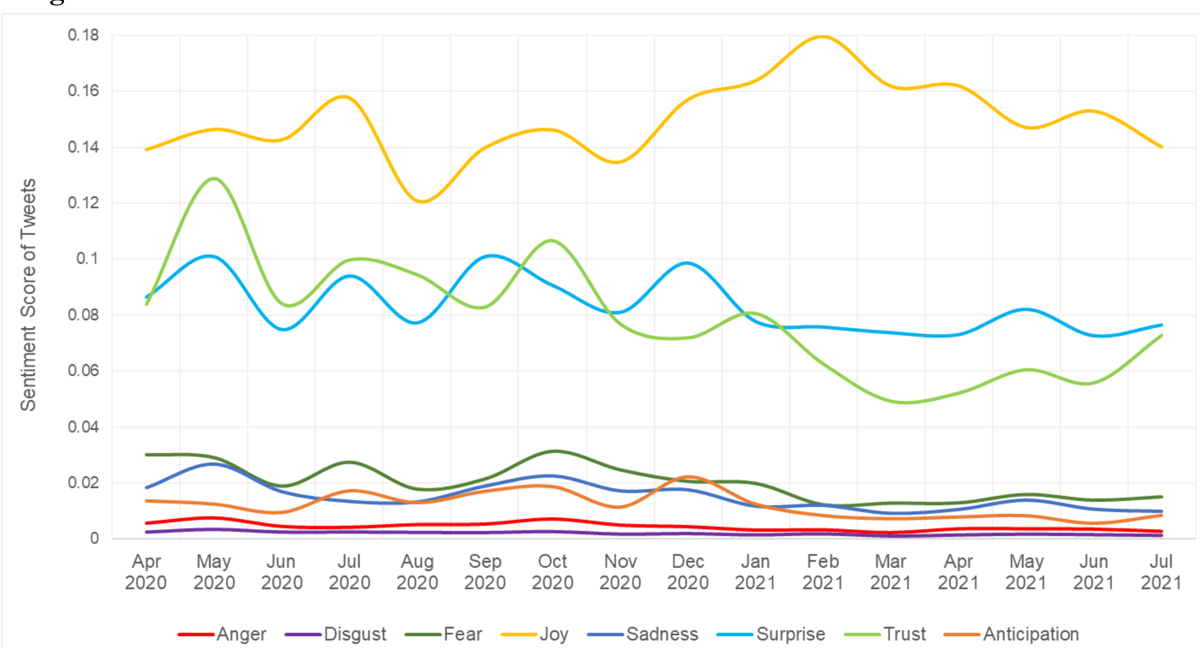
In addition to VADER sentiment classification, Plutchik's and POMS emotion profiles were also determined for the leaders' COVID-19 related tweets. As shown in Figure 19, joy (42.39%) and sadness (40.02%) emerged as the dominant emotions among the G7 leaders' tweets based on Plutchik's emotion classification. On the other hand, the proportion of tweets classified as anger was negligible at only 0.03%.

Figure 20. Relative Distribution of G7 Leaders' COVID-19 Tweets by Plutchik Sentiment

	Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise	Trust	Total
Boris Johnson	0.0%	0.6%	1.3%	1.0%	33.0%	45.4%	6.6%	12.1%	100%
Charles Michel	0.0%	0.2%	0.5%	0.3%	31.9%	55.9%	7.3%	3.8%	100%
Emmanuel Macron	0.1%	0.2%	0.7%	2.3%	52.8%	27.8%	2.7%	13.4%	100%
Italian Presidency Council	0.1%	1.1%	0.7%	4.6%	51.3%	11.9%	7.9%	22.4%	100%
Joe Biden	0.0%	0.5%	0.1%	1.5%	34.7%	47.4%	3.9%	11.8%	100%
Justin Trudeau	0.0%	0.7%	0.7%	1.7%	53.9%	32.5%	2.7%	7.8%	100%
Ursula von der Leyen	0.0%	0.5%	0.6%	0.1%	34.1%	52.0%	9.6%	3.1%	100%
Yoshihide Suga	0.0%	0.0%	0.0%	0.2%	58.3%	37.1%	3.1%	1.3%	100%

Figure 20 breaks down the Plutchik's emotion classifications by G7 leader. Relative to other leaders, Boris Johnson's tweets had a higher prevalence of disgust. The Italian Presidency Council's tweets had a relatively higher prevalence of anticipation, fear, and trust, and a much lower prevalence of joy. Moreover, Yoshihide Suga had the highest proportion of tweets classified as joy. Among all leaders, Ursula von der Leyen's tweets were most associated with surprise.

Figure 21. Sentiment Trend of G7 Leaders' Tweets based on Plutchik's Emotion Classification



Among the different emotions, joy was the only emotion to experience an uptrend as shown in Figure 21. The level of joy in leaders' tweets increased from August 2020 through February 2021 and then decreased from February 2021 to July 2021, but still remained higher than 2020 levels. The upward trend of joy coincides to a degree with the upward trend of positive sentiment noted in Figure 16. The other emotions generally show a constant or decreasing trend. Trust and fear, for example, exhibit a downtrend across the 16 months covered.

Figure 22. Profile of Mood States (POMS) Classification of G7 Leaders' COVID-19 Tweets

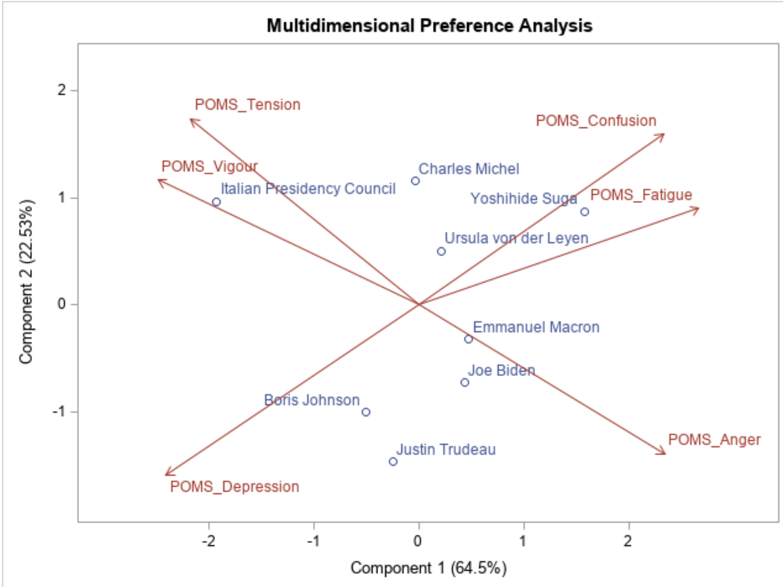
The POMS emotion classifications were also determined for the leaders' tweets. Unlike VADER and Plutchik's classification methods, POMS is based on mood states rather than hinged on a spectrum of positive and negative emotions. Based on the POMS emotion classification method, 65.2% of the leaders' tweets were categorized as exhibiting depression. On the other hand, only 3.2% of the tweets were classified as showing vigor. The notably high proportion of depression among the tweets perhaps stems from the nature of the model developed by Colneric and Demsar (2020) that was used to classify the POMS emotions. To illustrate, 40% of the tweets in the dataset used by the authors to train the POMS classification model were classified as depression while less than 2% were vigor-related tweets. This distribution was reflective of actual Twitter usage in that people tweet more about depression than they do about vigor. The results in Figure 22 show a similar distribution which suggests that the POMS method has a high tendency to classify tweets as depression-related.

Figure 23. Relative Distribution of G7 Leaders' COVID-19 Tweets by POMS Sentiment

	Anger	Confusion	Depression	Fatigue	Tension	Vigour	Total
Boris Johnson	10%	5%	67%	9%	5%	4%	100%
Charles Michel	10%	7%	60%	11%	6%	5%	100%
Emmanuel Macron	12%	6%	65%	9%	4%	3%	100%
Italian Presidency Council	6%	4%	72%	6%	7%	5%	100%
Joe Biden	11%	6%	65%	10%	5%	2%	100%
Justin Trudeau	11%	6%	67%	9%	4%	3%	100%
Ursula von der Leyen	11%	8%	61%	11%	6%	3%	100%
Yoshihide Suga	10%	8%	58%	15%	4%	3%	100%

Figure 22 shows the distribution of POMS emotions by G7 leader. Compared to Plutchik's emotions, the POMS emotions are more even in distribution across the different leaders with the disparities being less evident. The Italian Presidency Council appears to have lower prevalence of anger and confusion but higher prevalence of tension in their tweets compared to the other leaders. To examine the relative differences in the distribution at a deeper level than Figure 23 can portray, a PCA biplot is presented in Figure 24 below.

Figure 24. PCA Biplot of G7 Leaders Towards POMS Emotions



The biplot above projects the POMS emotions and leaders in a two-dimensional space using PCA and represents each POMS emotion as a vector. The two components combined explain 87% of the variation in the data which suggests the plot can be relied upon. Looking at the POMS emotions alone, confusion and fatigue, as well as tension and vigor, seem to have similarities based on their close proximity. Moreover, the association between the leaders and emotions can be inferred by inspecting the perpendicular projection of the leaders' points to each emotion's vector and seeing which leaders gravitate most towards a certain emotion. For example, Boris Johnson and Justin Trudeau appear to have the highest loadings towards the vector for depression. For vigor and tension, the Italian presidency council has the highest loading, followed by Charles Michel. Similarly, Yoshihide Suga and Charles Michel were most associated with confusion and fatigue. Trend-wise, Figure 25 shows that the POMS emotions in the leaders' tweets did not change significantly over time.

Figure 25. Sentiment Trend of Tweets based on POMS Emotions

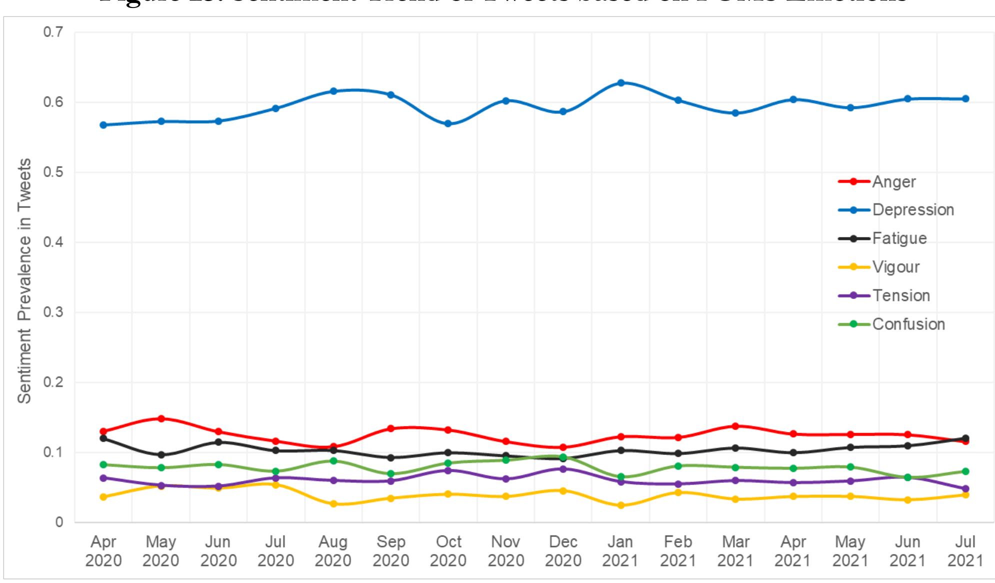
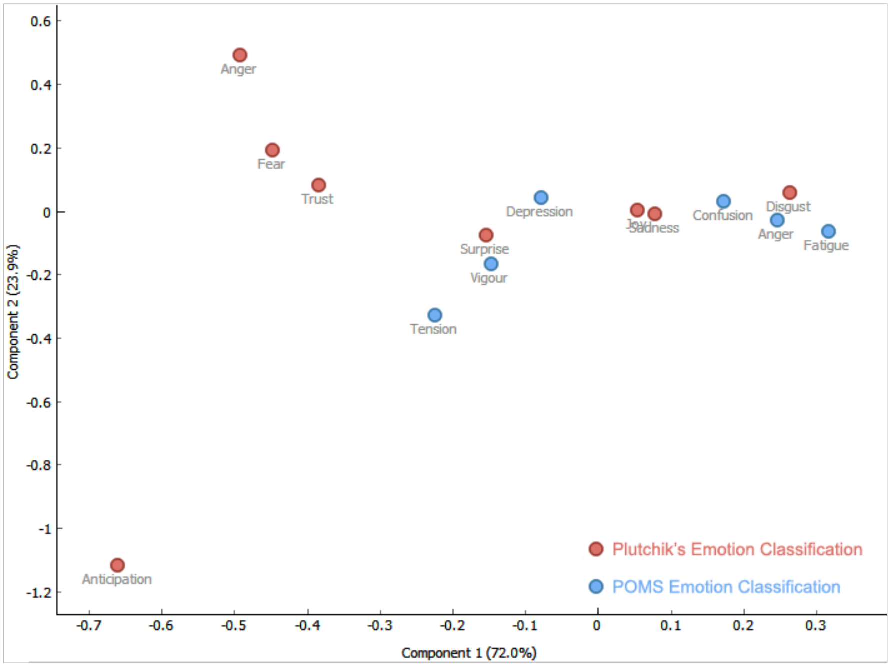


Figure 26. Correspondence Analysis of POMS and Plutchik's Emotions



The Plutchik's and POMS emotion classification methods resulted in different distributions of emotions. To see how the two classification methods are associated with each other in regards to the leaders' tweets, Figure 26 shows a correspondence analysis of the two categories of emotions. The Plutchik's emotions appear more spread out across the two-dimensional space while the POMS emotions are positioned closer to each other. This possibly suggests that Plutchik's emotions are more distinct from each other. All the six POMS emotions cover an area that roughly overlaps just four of Plutchik's emotions (surprise, joy, sadness, and disgust). In this way, the other four of Plutchik's emotions (anticipation, fear, trust, and anger) possibly explain aspects of emotion not covered by the six POMS emotions. Also apparent is that the emotion – anticipation - is most distant to other emotions of both categories. POMS emotions such as confusion, anger, and fatigue are similar to each other as well as to the Plutchik's disgust emotion. Similarly, POMS' vigor emotion and Plutchik's surprise emotion are closely positioned to each other.

Figure 27. G7 Leaders' Sentiments Regarding Selected COVID-19 Keywords

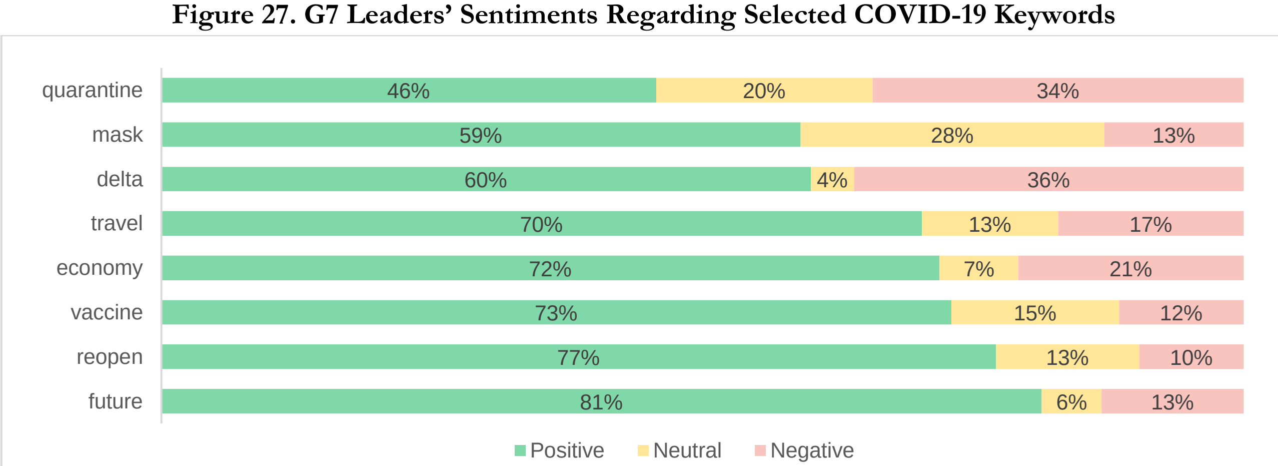


Figure 27 shows the overall sentiment of G7 leaders' regarding selected COVID-19 related keywords. Negative sentiments were strongest for the keywords "delta" (36%), which refers to the delta variant of the virus, followed by "quarantine" (34%). On the other hand, leaders' sentiments regarding the words "future", "reopen", "economy", "travel", "vaccine" were mostly positive. Among the keywords, "mask" was the most neutral word in terms of leaders' sentiments.

Figure 28. Trend of Leaders' Sentiment Towards Selected COVID-19 Keywords

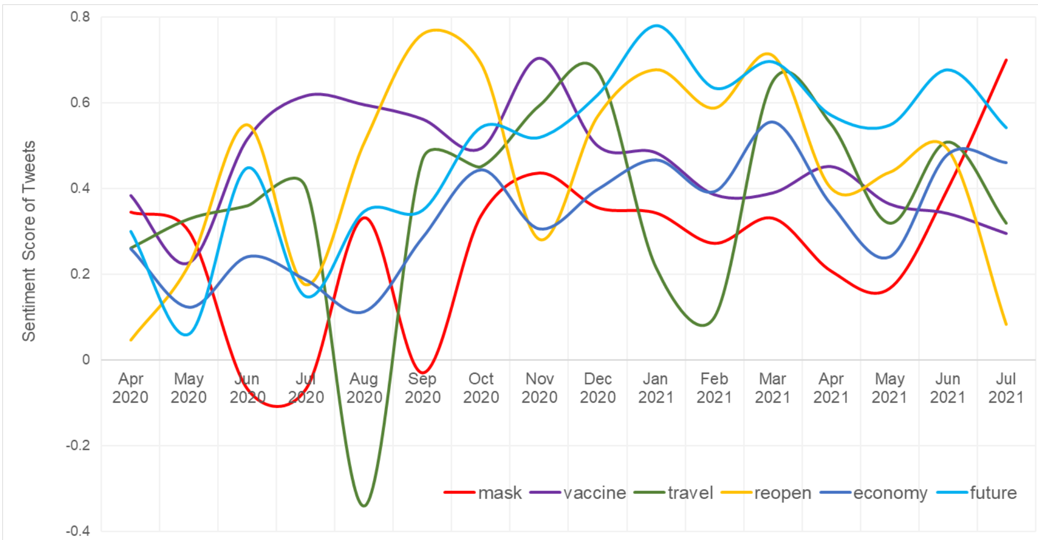


Figure 28 shows the trend of leaders' sentiments towards selected COVID-19 related keywords. Immediately apparent from the visualization is a large dip in the sentiment polarity regarding "travel" in August 2020. During this time, several countries imposed heightened travel and border restrictions amidst the rising cases all over the world. For instance, the European Union (EU) recommended halting travel from the United States in August 2020 due to the rising cases in the country (Kantis & Kiernan, 2021). It can also be observed that the leaders' sentiment polarity regarding "vaccine" was relatively low early in the pandemic and then peaked in November 2020 – the month when both Pfizer and Moderna reported their vaccines to be 95% effective (Ghosh, 2021). Leaders' sentiments regarding the keywords "future" and "economy" became increasingly positive as the pandemic progressed.

Table 8. Correlations between G7 Country-level Indicators and Polarity

Country-level COVID & Development Indicators	Correlation with Sentiment Polarity	Sig.
cty_new_cases_per_mil	0.020	0.050
cty_reproduction_rate	-0.009	0.381
cty_new_vaccinations	0.079	0.000
cty_fully_vaccinated_per_hundred	0.044	0.000
cty_stringency_index	-0.036	0.000
cty_covid_related_anxiety	-0.041	0.000
cty_population_density	0.067	0.000
cty_gdp_per_capita	-0.108	0.000
cty_hosp_beds_per_thousand	0.051	0.000
cty_human_development_index	-0.001	0.897
cty_social_support_index	-0.022	0.030
cty_corruption_index	-0.144	0.000
cty_happiness_index	0.009	0.387
cty_GINI_index	-0.158	0.000

Figure 28 revealed that the leaders' sentiments regarding key COVID-19 related keywords fluctuated as the pandemic progressed. This suggests a possibility that the leaders' sentiments may have been influenced by the development of COVID-19 in their respective countries or by other external country-specific factors. To investigate this, Table 8 shows the pairwise correlations of sentiment polarity with relevant COVID-19 related variables and country-level indicators. The descriptive statistics for these country-level variables can be found in Appendix D. Statistically significant correlations were noted for all variables except reproduction rate, human development index, and happiness index. While statistically significant, the magnitude of all the correlations is very small which is expected since tweet sentiment is primarily driven by the factors such as context, tweet content, word choice, language, tone, and secondarily by external factors. From the table, it can be inferred that leaders of countries with higher corruption index and GINI index (i.e. income inequality) were associated with lower sentiment polarity. Moreover, stringency index – which measures the level of COVID-19 related restrictions in a country – was negatively associated with sentiment polarity. This suggests that the leaders' sentiments, to an extent, became more negative as the level of restrictions became more stringent in their respective countries. On the other hand, the sentiment polarity of leaders was positively associated with the number of new vaccinations, fully vaccinated individuals, and hospital beds available in their respective countries.

Figure 29. Trend of COVID-19 Variables Relative to Sentiment Polarity

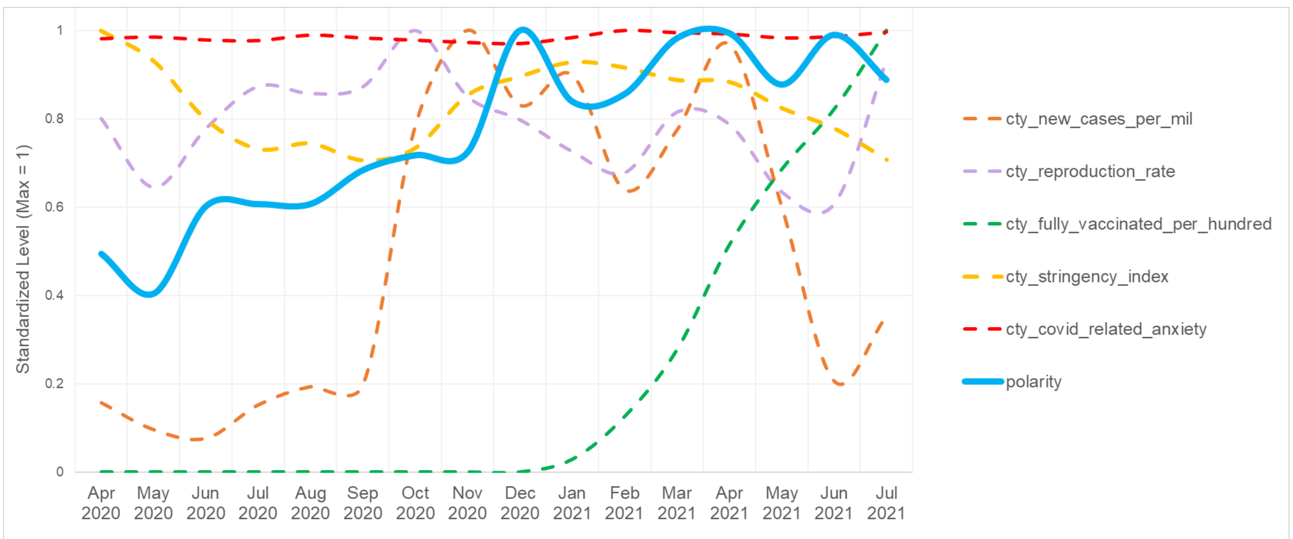
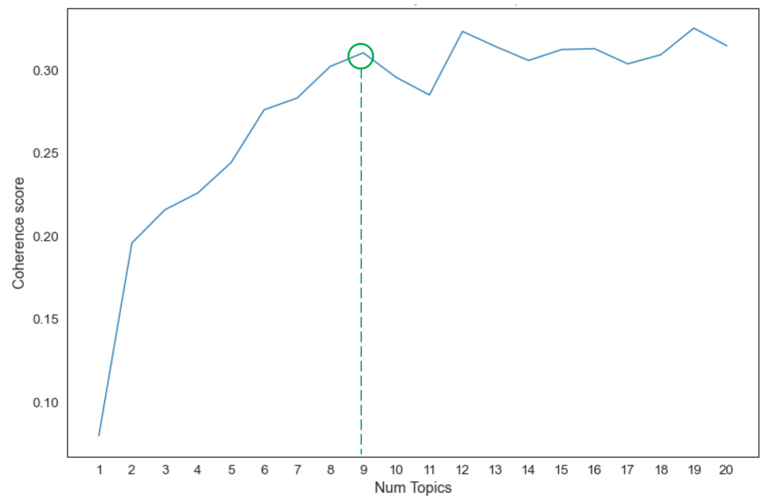


Figure 29 puts in perspective the trend of leaders' sentiment polarity relative to the trend of COVID-19 related variables. At a high level, the leaders' sentiment polarity increased and became more positive in 2020 and then stabilized in 2021. It can be also be seen that vaccination steeply increased beginning in 2021. While the polarity remained fairly constant during this time, it increased heading into January 2021 perhaps in anticipation of vaccination rollout and recovery. Moreover, the reproduction rate of the virus appears to have similar oscillations with the number of new cases, which is expected. The reproduction rate peaked in October 2020 and then sharply declined over the next several months. During this same time, the sentiment polarity of the leaders became more positive and remained constant thereafter.

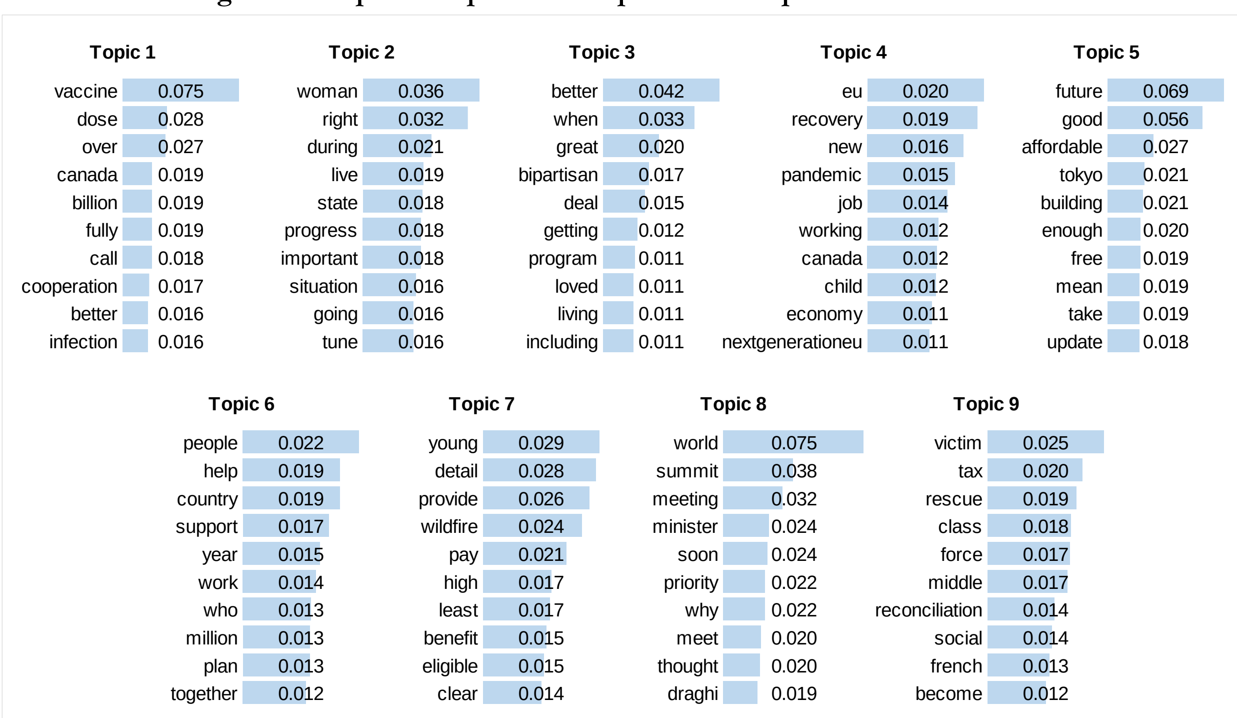
4.3. Topic Discovery

Figure 30. Coherence Score by Number of LDA Topics



In this subsection, the emergent topics identified using LDA topic modeling are summarized. Nine topics emerged from the G7 leaders' tweets regarding COVID-19. Based on the coherence score, the optimal number of topics (k) for the LDA topic model was determined to be nine (9). As can be seen in Figure 30, changes in the coherence score were negligible for models with greater than nine topics. Figure 31 shows the top words for each of the nine topics based on the per-topic-per-word probabilities produced by the final LDA model.

Figure 31. Top Words per LDA Topic based on per-word Probabilities



Based on the prominent words per topic shown in Figure 31, it can be inferred that the G7 leaders' discussion on Twitter regarding COVID-19 loosely revolved around nine main topics, namely: (1) vaccination, (2) situation monitoring and progress, (3) negotiations, (4) pandemic recovery, (5) rebuilding future, (6) community coordination, (7) crisis support and subsidies, (8) diplomacy, and (9) global news. The emergence of topics such as pandemic recovery (topic 4), rebuilding future (topic 5), and community coordination (topic 6) validates the potential topics hinted by the word co-occurrence network diagram (Figure 10) presented in the preliminary descriptive analysis. Figure 32 shows the word clouds per topic to provide a more detailed representation of the words within each topic.

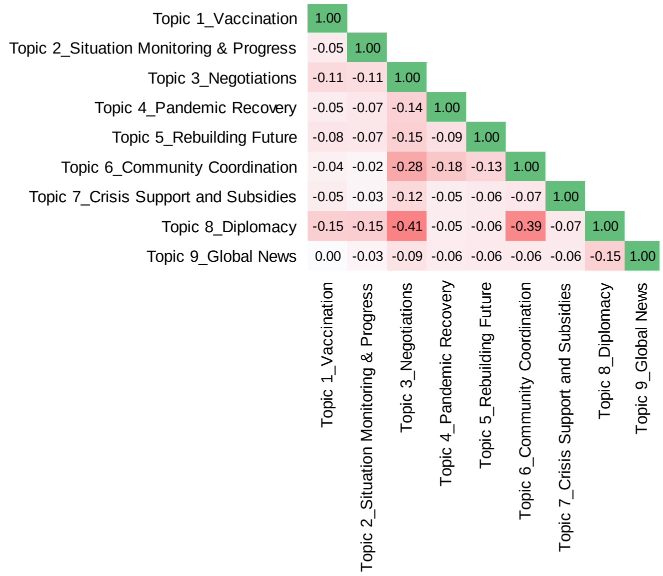
Figure 32. Word Cloud Representations of LDA Topics

Topic 1 involves discussion related to vaccination as evidenced by prominent keywords such as "vaccine", "dose", "fully", "infection", and "prevention". Topic 2 focuses on the leaders' discussion towards logistics of the pandemic situation and the current state of things regarding COVID-19. To illustrate, underlying keywords for Topic 2 include "situation", "progress", "information", "tune", "discuss", and "important". Topic 3 exudes a more political theme about the G7 leaders' negotiations and governance regarding COVID-19 as suggested by keywords like "bipartisan", "program", "administration", and "deal". Topic 4 captures the leaders' motivations to fight the pandemic and devise solutions to recover from it as hinted by keywords such as "recovery", "working", "fight", "pandemic", and "economy". Topic 5 revolves around rebuilding the future and is concentrated on current as well as post-pandemic issues and events. Sample prominent keywords for Topic 5 include "future", "building", and "preparedness". Topic 6 is associated with the leaders' intent to stimulate and maintain a sense of community and coordination among citizens during the pandemic as evidenced by keywords such as "people", "together", "community", and "support". Topic 7 is related to leaders' tangible support and subsidy provision towards not only COVID-19 but also other co-occurring crises such as forest wildfires as depicted by words such as "provide", "pay", "benefit", and "subsidy". Similar to Topic 3, Topic 8 reflects a political theme related to the G7 leaders' global diplomacy and meeting obligations. Top words for Topic 8 include "summit", "conference", "meet", "minister", and "world". Finally, Topic 9 is collectively associated with news and information about various global events discussed by leaders in the context of the COVID-19 pandemic as portrayed by words such as "afghanistan", "olympics", "incident", and "rescue". Since LDA provides a probability distribution across topics, a tweet can have elements of several topics and may not be exclusively about a single topic despite being classified according to its dominant topic (i.e. topic with the highest probability). Table 9 presents sample tweets for each topic.

Table 9. Sample Tweet per LDA Topic

Topic	Sample Tweet (Translated)	Sentiment	G7 Leader	Date
Topic 1_Vaccination	*@MichaelBuble @wick_22 @clarahughes @thisisourshotca Thanks for encouraging everyone to get vaccinated, Michael. It's such an important message. #ThisIsOurShotCA*	Positive	Justin Trudeau	04/28/2021 17:25
Topic 2_Situation Monitoring & Progress	*Thank you to the NHS and all of our critical workers for all you are doing to fight #coronavirus. You really are an inspiration. #ClapForOurCarers #ThankYouThursday #ClapForKeyWorkers*	Positive	Boris Johnson	04/02/2020 19:29
Topic 3_Negotiations	*I have written to all leaders of opposition parties to invite them to work together at this moment of national emergency.*	Positive	Boris Johnson	04/04/2020 8:48
Topic 4_Pandemic Recovery	*Our priority is to ensure that Europeans can travel freely in EU while limiting the spread of #COVID19. With various restrictions, it is now hard to know where we can travel & how. Today we propose to make it clearer, easier & safer, with a common colour code & set of measures.*	Positive	Ursula von der Leyen	09/04/2020 12:10
Topic 5_Rebuilding Future	*#ConferenzaStampa, Draghi: Approval for the gradual reopening of the country with calculated risk. This is largely the result of #campagnavaccinale*	Positive	Italian Presidency Council	05/20/2021 18:37
Topic 6_Community Coordination	*Community spirit is everywhere in Europe!* *Flag of Malta In Valetta, Josephine sews masks for her neighbours. Thanks to her dexterity, old shirts have a second life and help protect the community.* *Together, we will overcome the pandemic.*	Positive	Ursula von der Leyen	12/31/2020 12:07
Topic 7_Crisis Support and Subsidies	*We've created a number of programs and benefits to help you make ends meet - there's the wage subsidy, the rent subsidy, the Recovery Benefit, the Sickness Benefit, the Caregiving Benefit, and more.*	Positive	Justin Trudeau	01/25/2021 23:01
Topic 8_Diplomacy	*Solidarity must underpin our global response to #COVID19.* *Not a solidarity based on a misplaced clear conscience. Not a solidarity of words.* *But a solidarity of action. Tangible and real. To forge a more caring society for the future. #OACPSSummit2020*	Positive	Charles Michel	06/03/2020 2:01
Topic 9_Global News	*With the pandemic, Lourdes could welcome its pilgrims and tourists for several months. Its shops, its economy, were tested. 124 million of aid was earmarked for Lourdais tourism. For its territories, the state still meets this!*	Positive	Emmanuel Macron	07/16/2021 18:31

Figure 33. Correlation Matrix of LDA Topics



Highly related topics are undesirable because they signal ineffective differentiation of tweet content into latent groups and imply that topics may be further merged into a single topic. The correlation matrix of topics shown in Figure 33 suggests that the identified topics are generally distinct from each other since all the pairwise correlations are negative and small in magnitude. The largest correlation is -0.41 which reflects a moderate negative correlation between Topic 8 (diplomacy) and Topic 3 (negotiations). The UMAP projection in Figure 34 also reveals that Topic 8 and Topic 3 are most distant to one another based on relative probability distributions across all the tweets. This suggests that the two topics are most dissimilar. Furthermore, the UMAP projection shows that Topics 1,2,5,6, and 7 are most similar to each other. Upon close inspection, all these five topics seem to be most closely related to COVID-19 at the core as they all describe tangible, solution-oriented efforts of the leaders with respect to handling, controlling, and recovering from the COVID-19 pandemic.

Figure 34. UMAP Projection of LDA Topics

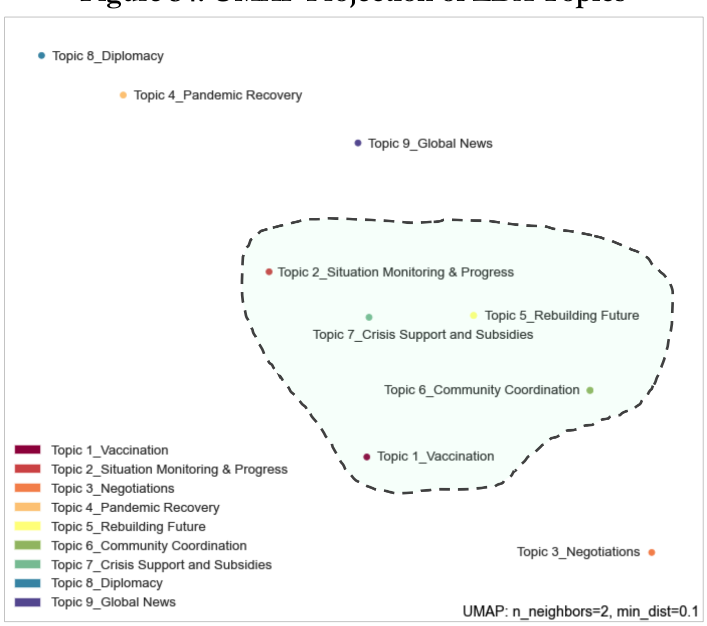


Figure 35. G7 Leaders' Sentiments Regarding LDA Topics

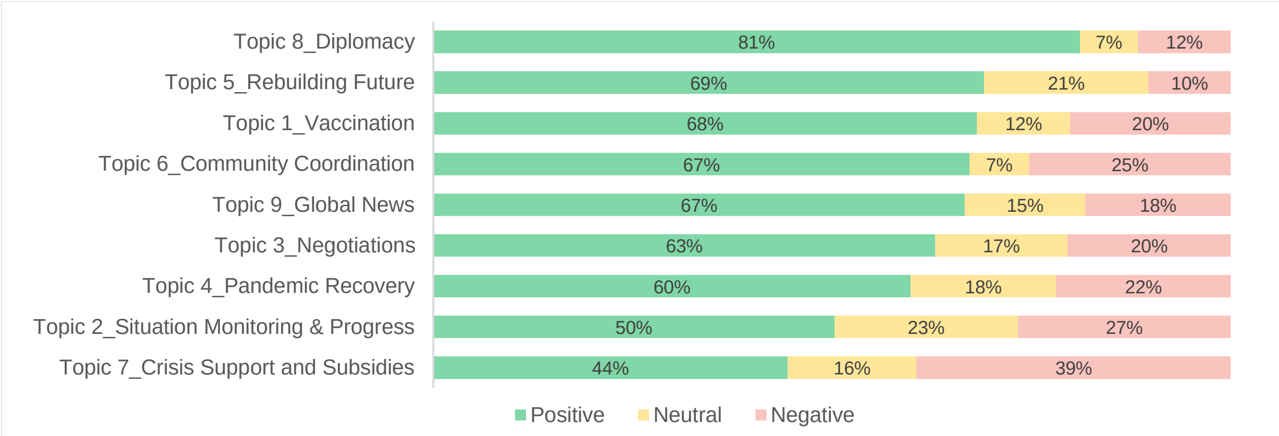
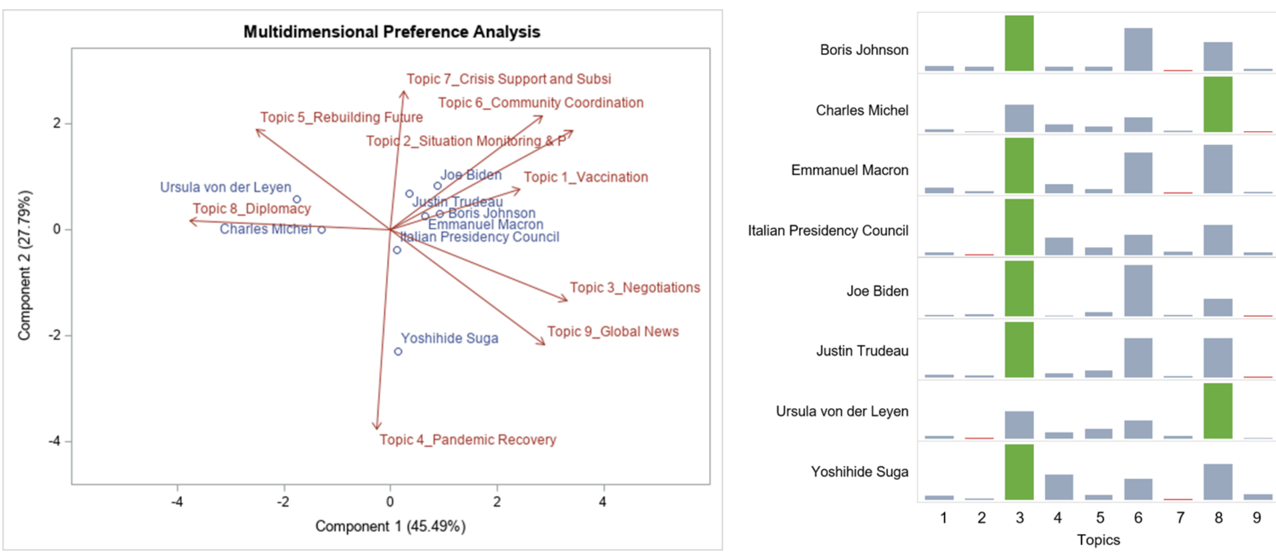


Figure 35 shows the sentiment distribution per topic. Leaders were most positive regarding Topic 8 (diplomacy) – 81%, followed by Topic 5 (rebuilding future) – 69%, Topic 1 (vaccination) – 68%, and Topic 6 (community coordination) – 67%. On the contrary, leaders' tweets related to Topic 7 (crisis support and subsidies) were the least positive (44%) and most negative (39%).

Figure 36. PCA Biplot and Barplot of G7 Leaders' Affinity towards LDA Topics



As shown in Figure 36, some G7 leaders were more associated with certain topics. For instance, the PCA biplot suggests that Ursula von der Leyen and Charles Michel talked about Topic 8 (diplomacy) more relative to other leaders. This is expected given both these G7 leaders are representatives of the European Union (EU). Also apparent from the biplot is that Yoshihide Suga was most associated with Topic 4 (pandemic recovery). Moreover, the Italian Presidency Council's tweets were most well-rounded across the topics as suggested by its location in the PCA biplot which is closest to the origin. Relative to the other leaders, Joe Biden tweeted more about Topic 6 (community coordination). For all the leaders, Topic 3 (negotiations) and Topic 8 (diplomacy) had the highest prevalence in the tweets.

Figure 37. G7 Leaders' Sentiment Polarity Towards LDA Topics

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9
Boris Johnson	0.40	0.38	0.39	0.25	0.52	0.47	0.11	0.56	0.48
Charles Michel	0.65	0.17	0.27	0.53	0.22	0.23	0.14	0.48	0.35
Emmanuel Macron	0.30	-0.12	0.26	0.10	0.33	0.30	0.08	0.33	0.28
Italian Presidency Council	0.13	0.23	0.06	0.10	0.36	0.14	-0.02	0.31	0.14
Joe Biden	0.37	0.01	0.23	0.01	0.24	0.21	-0.05	0.34	0.30
Justin Trudeau	0.34	0.28	0.40	0.35	0.38	0.38	0.17	0.50	0.48
Ursula von der Leyen	0.34	0.08	0.36	0.63	0.45	0.23	0.34	0.56	0.38
Yoshihide Suga		0.80	0.32	0.37	0.16	0.42	-0.65	0.58	0.45

In terms of sentiment, Figure 37 suggests that Yoshihide Suga was most positive regarding Topic 2 (situation monitoring and progress) and most negative towards Topic 7 (crisis support and subsidies). Among the leaders, only Emmanuel Macron held negative sentiments towards Topic 2 (situation monitoring and progress). It can also be seen that Charles Michel was the most positive towards Topic 1 (vaccination). Figure 38 projects the leaders into a two-dimensional space based on their LDA topic variation using t-SNE, MDS, Isomap, and UMAP. All four methods show three clusters of leaders in terms of topic similarity: (1) Yoshihide Suga and the Italian Presidency Council, (2) Charles Michel and Ursula von der Leyen, and (3) Joe Biden, Boris Johnson, Emmanuel Macron, and Justin Trudeau.

Figure 38. Projection of G7 Leaders based on LDA Topics using Manifold Learning Methods

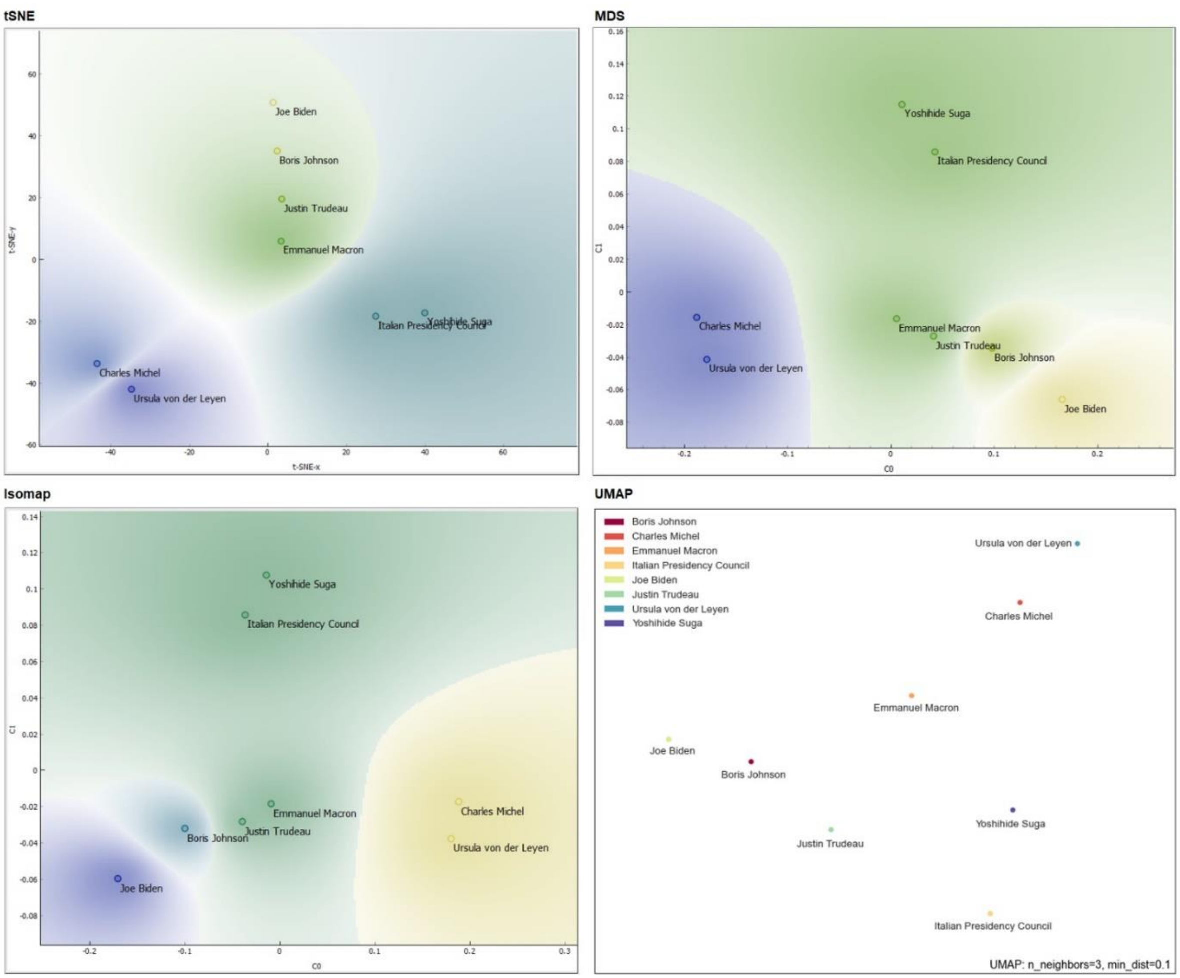


Figure 39. UMAP projection of G7 Leaders based on LDA Topic Variation by 4-month Periods

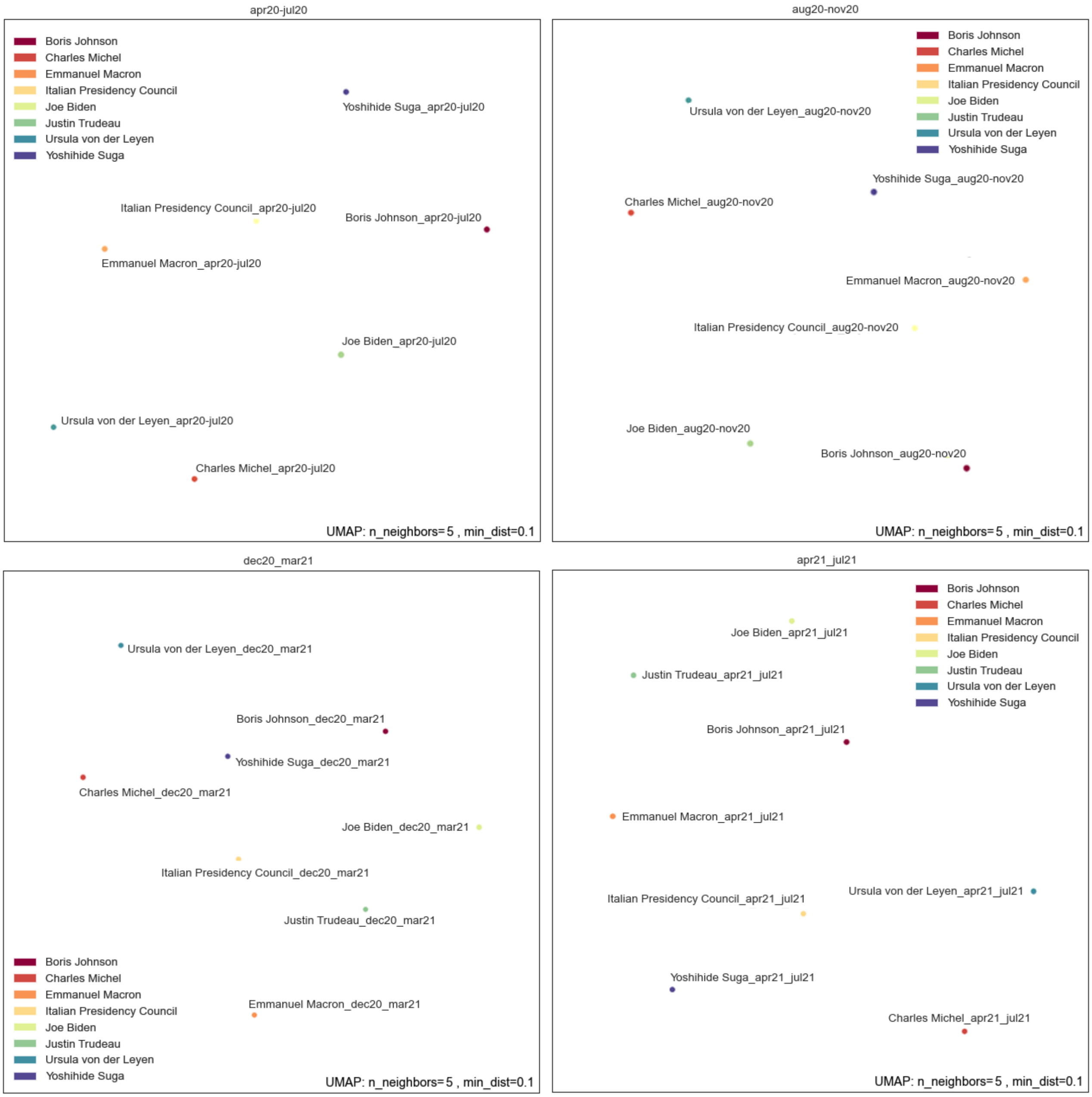


Figure 39 shows the UMAP projection of leaders based on LDA topic variation by 4-month periods from April 2020 to July 2021. Shifts in the location of the leaders' points throughout the four UMAP plots indicate changes in the leaders' topic distribution over time. For instance, the point referring to Boris Johnson appears to be plotted approximately in the same location throughout all periods except August 2020 to November 2020. This suggests that Boris Johnson's discussion on Twitter regarding COVID-19 topics deviated from usual in August 2020 to November 2020, but was otherwise consistent throughout the 16 months. Similarly, Yoshihide Suga's points appear to be plotted in the same locations for the first three time-periods from April 2020 to March 2021, but relatively distant in the last time-period from April 2021 to July 2021.

Figure 40. LDA Topics by Polarity (x), Virality (y), and Frequency (Bubble Size)

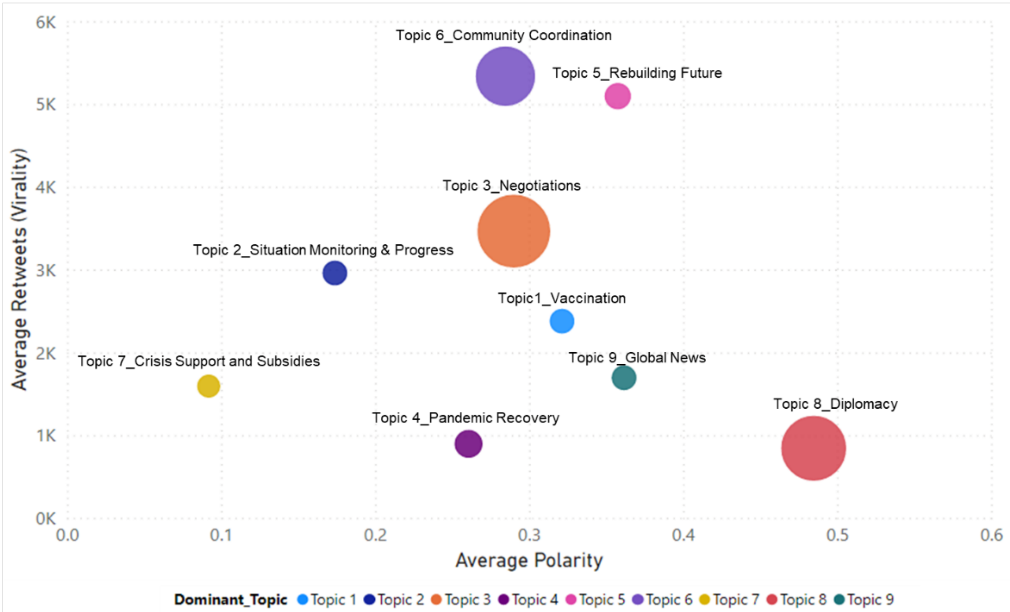


Figure 40 shows a comparison of the nine topics based on sentiment polarity, virality, and frequency or prevalence. The size of the bubbles reveals that Topic 3 (negotiations), Topic 8 (diplomacy), and Topic 6 (community coordination) were most prevalent in the leaders' tweets. The most viral topics were Topic 6 (community coordination) and Topic 5 (rebuilding future) based on retweets. Moreover, Topic 8 (diplomacy) was the most positive in terms of sentiment but the least viral in terms of the number of retweets.

Figure 41. Relative Distribution of LDA Topics by Plutchik's Emotion Classification

	Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise	Trust	Total
Topic 1_Vaccination	0.00%	1.3%	0.6%	1.9%	42.2%	31.2%	4.5%	18.2%	100%
Topic 2_Situation Monitoring & Progress	0.00%	1.2%	0.6%	4.3%	38.5%	36.0%	8.7%	10.6%	100%
Topic 3_Negotiations	0.03%	0.5%	0.7%	1.8%	43.7%	37.5%	4.2%	11.5%	100%
Topic 4_Pandemic Recovery	0.00%	0.0%	0.4%	2.2%	51.8%	25.2%	6.9%	13.5%	100%
Topic 5_Rebuilding Future	0.00%	0.0%	0.0%	4.9%	45.7%	26.0%	8.1%	15.2%	100%
Topic 6_Community Coordination	0.04%	0.8%	0.3%	1.7%	37.3%	45.4%	4.1%	10.3%	100%
Topic 7_Crisis Support and Subsidies	0.00%	0.0%	1.0%	3.0%	45.5%	32.3%	8.1%	10.1%	100%
Topic 8_Diplomacy	0.04%	0.3%	0.5%	0.4%	43.7%	42.5%	6.5%	6.0%	100%
Topic 9_Global News	0.00%	1.3%	0.0%	0.6%	44.0%	38.4%	5.0%	10.7%	100%

Figure 41 shows the Plutchik's emotion profile for each of the nine identified topics. Among all topics, Topic 1 (vaccination) was most associated with trust (18.2%). This suggests that the G7 leaders displayed trust in tweets about vaccination more than any other topic. Topic 5 (rebuilding future) had a relatively higher proportion of fear (4.9%) and surprise (8.1%). Moreover, Topic 7 (crisis support and subsidies) had the highest proportion of disgust (1.0%) relative to other topics. Finally, Topic 1 (vaccination), Topic 2 (situation monitoring and progress), and Topic 9 (global news) were most associated with anticipation.

Figure 42. Trend of LDA Topic Prevalence in G7 Leaders' Tweets

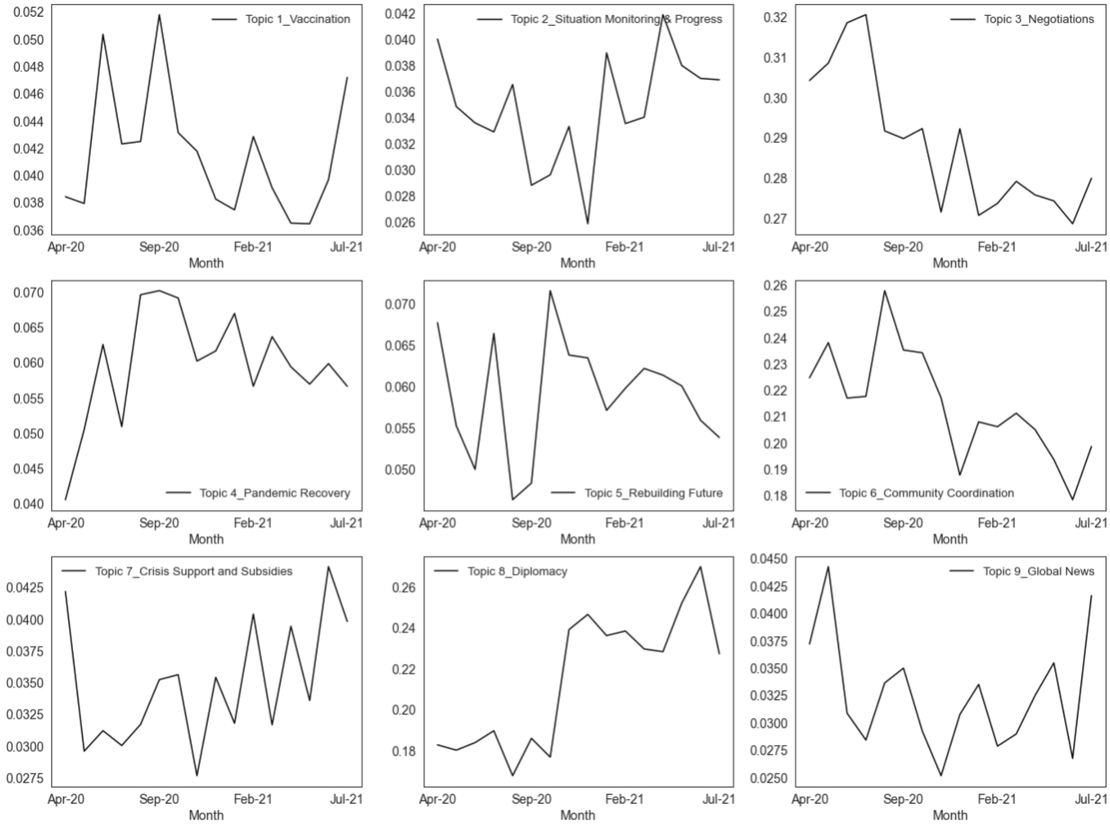
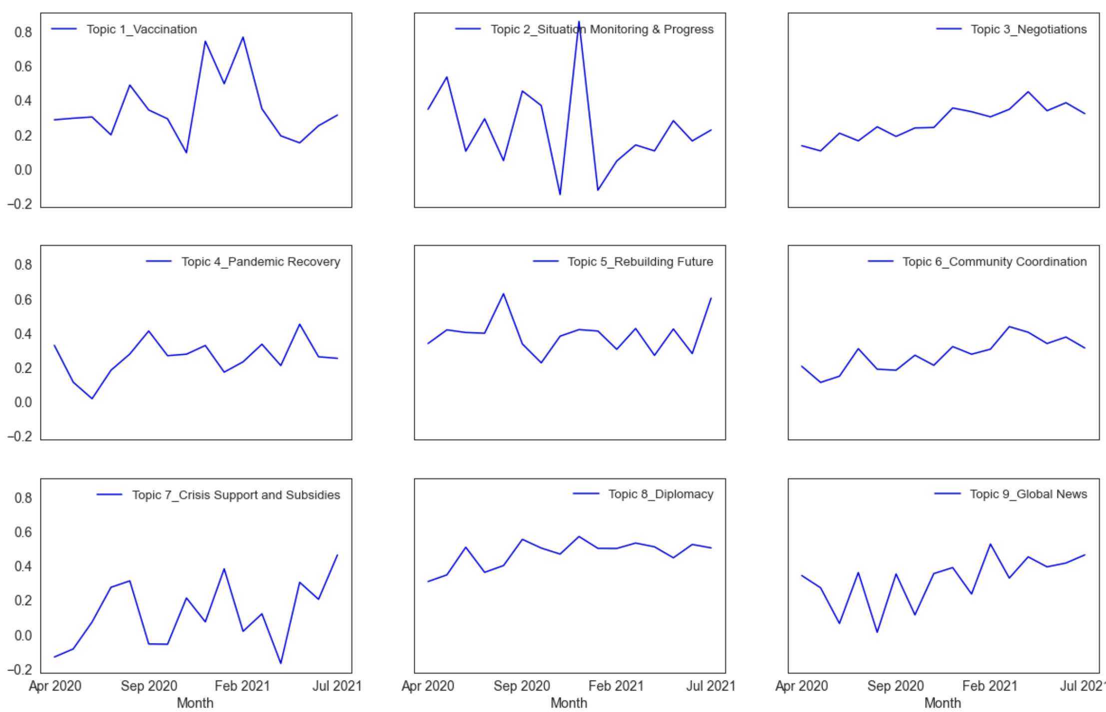


Figure 43. Sentiment Polarity Trend of LDA Topics



Figures 42 and 43 show the prevalence and sentiment trend of topics from April 2020 to July 2021. Across the duration covered, the prevalence of Topic 6 (community coordination) and Topic 3

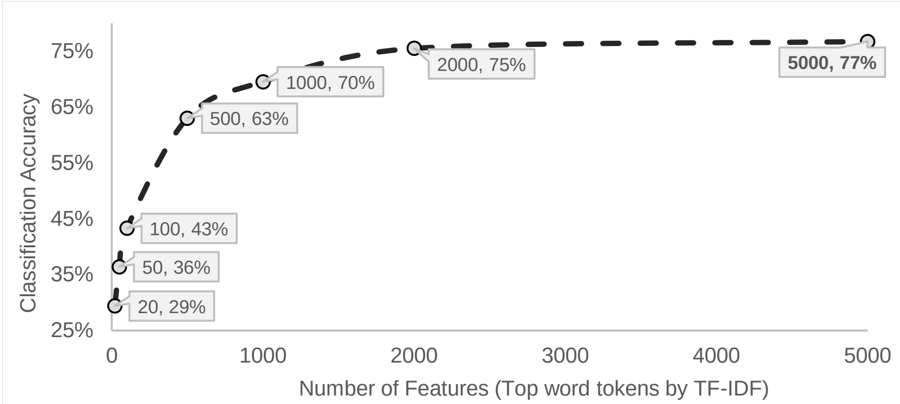
(Negotiations) decreased over time while the prevalence of Topic 7 (crisis support and subsidies) and Topic 8 (diplomacy) notably increased. The prevalence of Topic 2 (community coordination) fluctuated periodically and was high early in the pandemic (April 2020); it then decreased as the year 2020 passed and then increased again in 2021. Moreover, significant oscillations are apparent in the prevalence of Topic 2 (community coordination) and Topic 7 (crisis support and subsidies). These fluctuations may be reflective of sudden surges or waves of COVID-19 cases to which leaders reacted by tweeting more about these topics. Also noteworthy is that Topic 1 (vaccination) was more prevalent in 2020 perhaps in anticipation of the vaccination programs compared to 2021 when the actual rollout of vaccination commenced. In Figure 43, the trend of leaders' sentiment polarity regarding the topics is shown. Sentiments regarding topics 3,4,5,6, and 8 remained fairly constant throughout the 16 months covered. It can be seen that the sentiment polarity towards Topic 1 (vaccination) increased sharply from around December 2020 until March 2021. During this time, mass vaccination programs were being rolled out globally which might explain why leaders' sentiments were more positive regarding vaccination.

Table 10. Performance Metrics of LDA Topic Classification Models

Model (5000 features)	Accuracy	Precision	Recall	F1	AUC
Ensemble (Stack)	0.770	0.770	0.773	0.770	0.949
Logistic Regression	0.761	0.763	0.768	0.761	0.957
Stochastic Gradient Descent	0.761	0.761	0.763	0.761	0.860
Multilayer Perceptron (MLP)	0.756	0.758	0.765	0.756	0.950
Linear SVM	0.663	0.662	0.663	0.663	0.921
Naive Bayes	0.531	0.560	0.716	0.531	0.910
kNN	0.120	0.096	0.692	0.120	0.725

To better understand and validate the predictive nature of the nine topics determined by the LDA model, seven machine learning algorithms were applied. In the field of machine learning, predictive modeling is often referred to as supervised learning. If the predicted variable is categorical, the prediction task at hand is a classification problem. For continuous target variables, regression models are more suited for the prediction task (Ye, 2013). Since the goal is to predict or classify the dominant topic of the leaders' tweets among the nine identified topics, classification models are appropriate. Several machine learning algorithms are used for multi-label text classification tasks. As different algorithms produce different results depending on the nature of the data, it is important to compare different techniques to unravel which algorithm performs for the data at hand. Using the top 5000 TF-IDF word vectors as input features, seven machine learning algorithms were used to train and test the data with 10-fold cross-validation. As can be seen in Table 10, the models performed well in predicting the topic of the G7 leaders' tweets. The ensemble model combining the other six constituent models had the highest accuracy of 77%, followed by Logistic Regression and Stochastic Gradient Descent (76.1)%, and MLP (75.6%). The precision, recall, F1-score, and AUC metrics for these models also indicate healthy model fit (See Appendix E for ROC Curves per topic class). On the other hand, the kNN algorithm performed poorly with an accuracy of only 12%.

Figure 44. Model Performance (Accuracy) by Number of Features



It was computationally possible to build the final predictive model using 5000 features with a resulting accuracy of 77%. However, Figure 44 shows that using only 1000 features (top words by TF-IDF) results in a fairly high accuracy of 70% for predicting the latent dominant topic of the G7 leaders' tweets. This implies that it is possible to conduct accurate topic classification with as little as 1000 features for this type of data if computational resources are limited.

Figure 45. Confusion Matrix of Recall Values by Topic Class (Ensemble)

Actual \ Predicted	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9
Topic 1	81.5 %	0.0 %	6.3 %	1.7 %	0.4 %	6.3 %	0.0 %	2.9 %	0.8 %
Topic 2	1.7 %	75.3 %	7.1 %	1.7 %	1.3 %	5.9 %	0.8 %	5.0 %	1.3 %
Topic 3	1.6 %	1.2 %	75.4 %	3.0 %	1.9 %	9.9 %	0.4 %	6.0 %	0.5 %
Topic 4	1.5 %	0.7 %	5.4 %	85.6 %	0.7 %	1.7 %	0.0 %	3.9 %	0.5 %
Topic 5	0.6 %	0.6 %	3.9 %	0.3 %	86.4 %	5.1 %	0.3 %	2.7 %	0.0 %
Topic 6	2.1 %	1.5 %	19.7 %	3.0 %	3.0 %	62.2 %	1.1 %	6.0 %	1.5 %
Topic 7	2.0 %	0.0 %	7.3 %	3.3 %	0.0 %	4.6 %	74.8 %	7.9 %	0.0 %
Topic 8	0.5 %	1.6 %	11.0 %	3.2 %	1.4 %	3.7 %	0.4 %	77.7 %	0.5 %
Topic 9	1.3 %	0.8 %	7.2 %	0.4 %	0.8 %	7.2 %	0.8 %	1.7 %	79.7 %

As can be seen in Figures 45 and 46, the best model (ensemble) was able to correctly classify all nine topics with decent accuracy, ranging from 62.2% for Topic 6 - community coordination (most difficult to classify) to 86.4% for Topic 5 – rebuilding future (easiest to classify).

Figure 46. LDA Topic Classification Chart for Ensemble Model

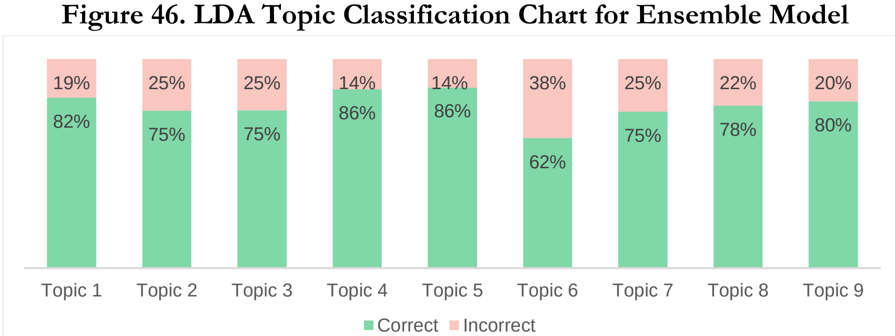
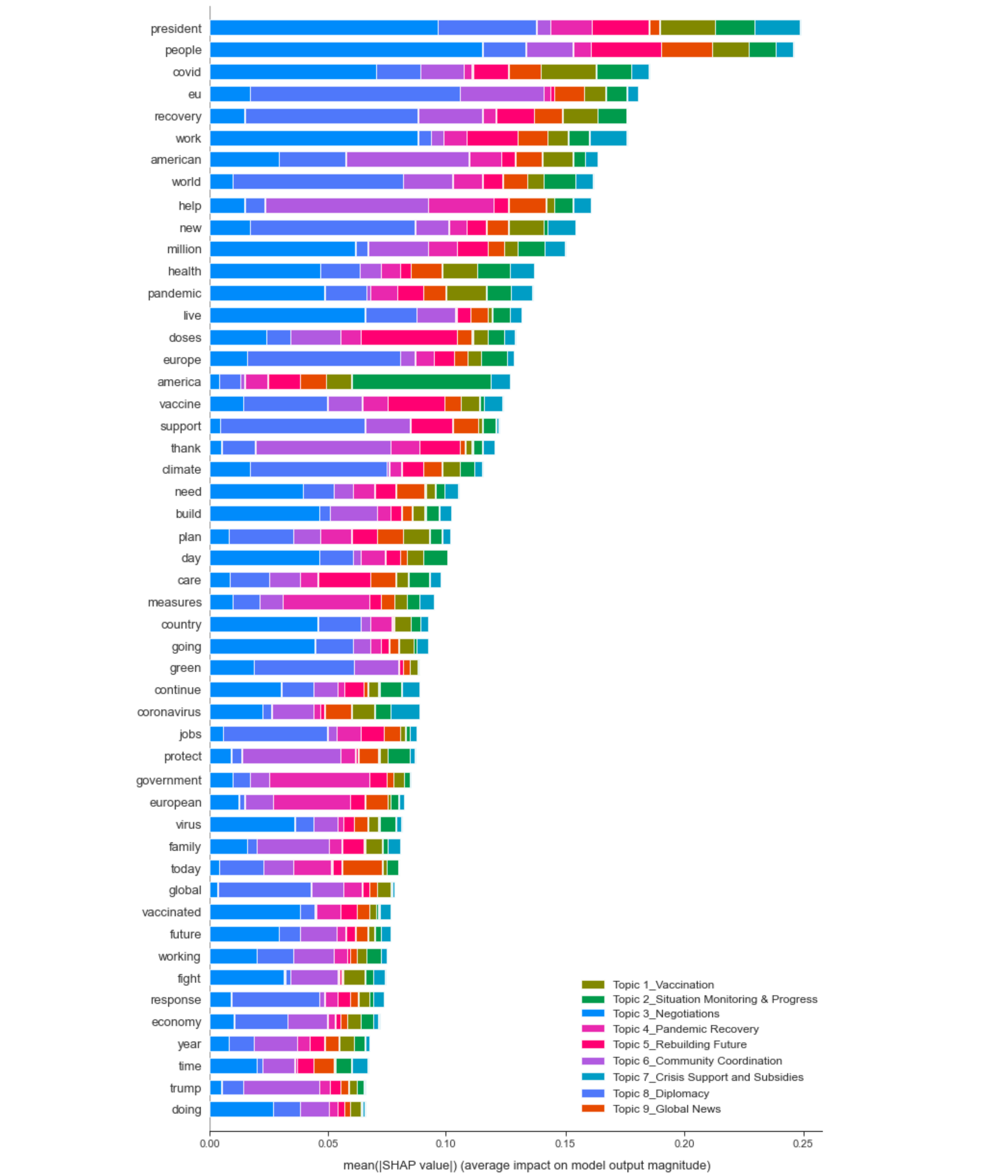
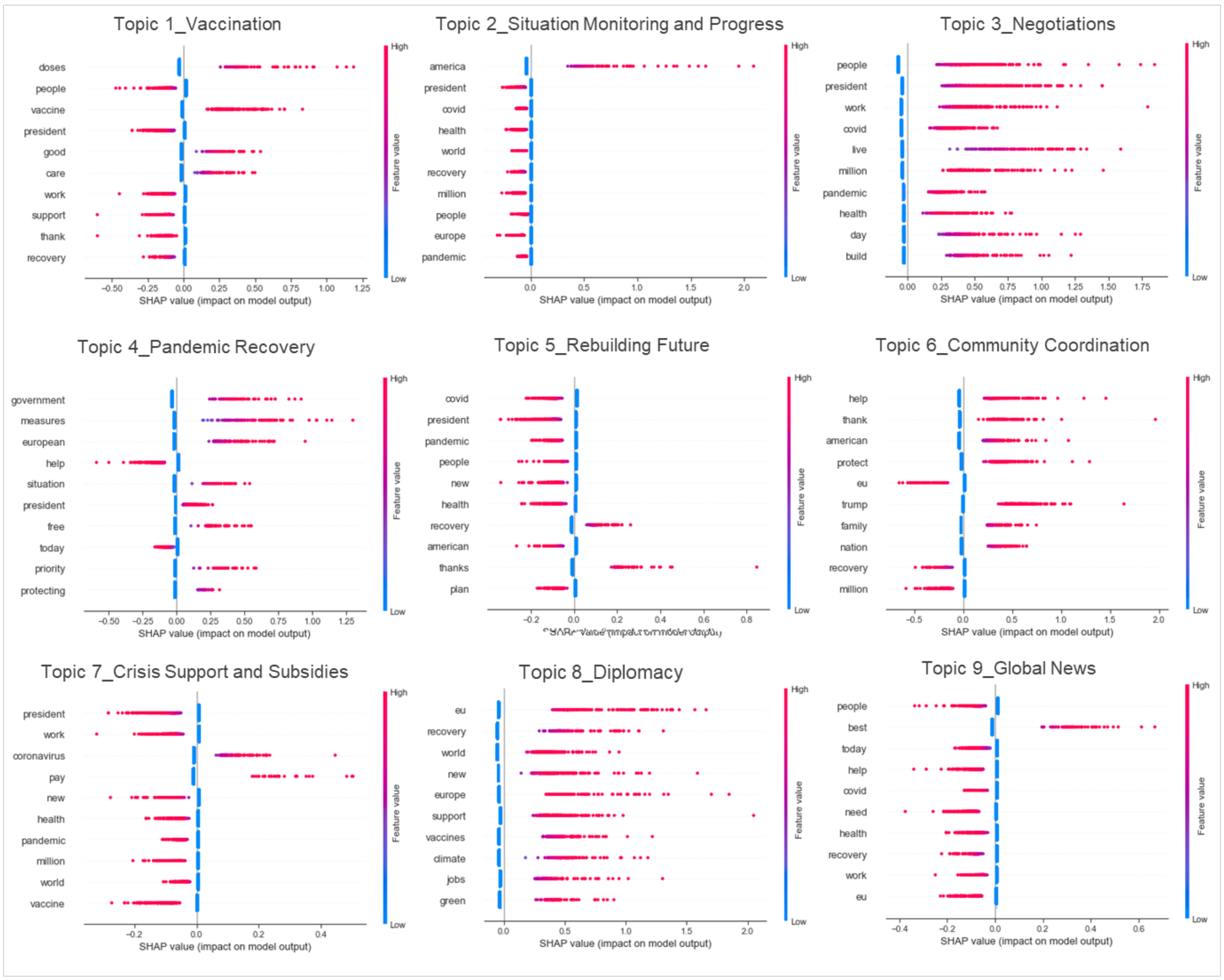


Figure 47. Top 50 Most Important Words for Topic Classification based on SHAP Values



In addition to model performance and accuracy, model interpretation and explanation form equally important parts of the overall assessment of a predictive model (Shikun, 2021). The game theory-based method - SHapley Additive exPlanations (SHAP) - was applied to explain the multi-class topic classification model. Using SHAP values, Figure 47 explains the predictive model by showing the top 50 most important words (i.e. features) based on their contribution towards classifying the nine identified topics. In terms of global feature importance, the top five most important words for differentiating the topics were "president", "people", "covid", "eu", and "recovery". To further interpret the predictive model, Figure 48 shows the SHAP summary plots for each topic class which illustrate the feature importance and feature effects of the top 10 words per topic

Figure 48. Top 10 Words per Topic Class based on SHAP Values



It can be observed that "government" and "measures" are the top two words with the highest and most positive impact towards classifying a tweet as Topic 4 (pandemic recovery). The high impact can be inferred from the red color while the positive direction is shown by the x-axis. Similarly, words such as "pay" and "coronavirus" contribute most towards classifying a tweet as Topic 7 (crisis support and subsidies). On the other hand, words like "president" and "work" contribute most in pushing the classification against Topic 7 as suggested by their negative SHAP values in the x-axis. For tweets classified as Topic 3 (negotiations), the most impactful words were "people", "president", "work", and "covid". Overall, the top 10 words per topic based on SHAP values generally coincide with the top 10 words per topic based on the per-word probabilities shown in Figure 31. These results validate that the proposed topic model performs well both in terms of classification accuracy as well as intuitive interpretability.

5. DISCUSSION AND CONCLUSIONS

Principal Findings. Social media platforms such as Twitter have emerged as powerful communication tools for political leaders to easily and effectively share information with the public during crises. The results have shown that the G7 leaders extensively leveraged Twitter as a tool to inform the public regarding COVID-19. Collectively, the eight selected G7 leaders posted 9,918 COVID-19-related tweets from April 2020 to July 2021. Justin Trudeau of Canada and Joe Biden of the United States tweeted most frequently, averaging 10.8 and 5.5 tweets per day respectively. The most common words that emerged from the leaders' tweets were "together", "people", "vaccine", "president", and "country". The leaders generally showed positive sentiments (69% positive, 12% neutral, 19% negative) regarding COVID-19. Among the leaders, Ursula von der Leyen's tweets were the most positive in terms of sentiment (81% positive) while Joe Biden and the Italian presidency Council used the least positive tone in their tweets. Trend analysis revealed an uptrend in the overall sentiment polarity across the duration covered and showed that leaders were more positive in 2021 compared to 2020. Boon-Itt and Skunkan (2020) also found that sentiment became increasingly positive as the pandemic progressed in conjunction with additional news being reported. Furthermore, oscillations in leaders' sentiments were found to be in part related to vaccination rollout, safety restrictions, and surges in COVID-19 cases. Small, but statistically significant correlations were also noted between tweet sentiment polarity and country-specific developmental variables such as COVID-19 stringency index, corruption index, and GINI index. Moreover, the top three prominent emotions that emerged from the leaders' tweets were joy (42%), sadness (40%), and trust (9.9%) according to Plutchik's emotion classification; and depression (65%), anger (10.5%), and fatigue (9.7%) according to the POMS emotion classification. It was observed that Plutchik's emotions varied more over time compared to POMS emotions. Leaders were more positive towards words such as "future" (81% positive) and "reopen" (77% positive) and more negative towards keywords like "quarantine" (34% negative) and "delta" (34% negative). Topic modeling using LDA uncovered that the G7 leaders discussed nine main topics regarding COVID-19: vaccination, situation monitoring and progress, negotiations, pandemic recovery, rebuilding future, community coordination, crisis support and subsidies, diplomacy, and global news. Compared to previous studies such as Xue et al. (2020) who analyzed tweets from March to April 2020, it was noted that topics relating to the origin of the virus in Wuhan, China were no longer as prevalent in the leaders' tweets after April 2020. In terms of sentiments, Topic 8 - diplomacy was the most positive (81%), followed by Topic 5 – rebuilding future (69% positive) and Topic 1- vaccination (68% positive). Related studies have also shown that sentiments regarding vaccination are generally positive. For example, Villavicencio et al., (2021) found that public sentiment on Twitter regarding vaccination was 83% positive. On the other hand, Topic 7 (crisis support and subsidies) was the most negative in terms of sentiment. Topic 1 (vaccination) was associated closely with trust and Topic 5 (rebuilding future) was discussed with higher levels of fear and surprise. The best predictive model (ensemble) was 77% accurate in classifying the G7 leaders' COVID-19 related tweets into the nine identified topics. Tweets about Topic 4 (pandemic recovery) and Topic 5 (rebuilding future) were the easiest to classify whereas tweets related to Topic 6 (community coordination) were the most difficult to classify. Model explanation based on SHAP values illuminated that "president", "people", "covid", "eu", and "recovery" were the five most important words for differentiating the topics.

Practical Policy Implications. During a public health crisis like COVID-19, world leaders are responsible for providing relevant information and reassuring the public. A study by Haman (2020) found that there was a significantly higher growth rate in the Twitter followers of world leaders during the COVID-19 pandemic compared to pre-pandemic levels. This suggests that netizens increasingly rely on world leaders and relevant health authorities for information during public health emergencies. In today's digital landscape, political leaders should increasingly take advantage of platforms like Twitter to provide swift, transparent, and relevant health information to their citizens. Moreover, the volume and frequency of tweets are indicative of how proactive a leader is during crisis situations. During pandemics, leaders must be proactive and compete for attention while conveying relevant messages to information-seeking citizens (Ratzan et al., 2020). Haman (2020) found that state leaders who posted more frequently on Twitter gained more followers. Based on the results, Justin Trudeau and Joe Biden are good examples among the G7 leaders in regards to posting proactively regarding COVID-19. Proactive health communication by authorities such as world leaders also helps shield against disinformation, such as false information regarding the effectiveness of vaccination. Indeed, this is supported by the finding that leaders showed more positive sentiments and a higher level of trust while discussing Topic 1 (vaccination). Consistent and proactive sharing of information also helps establish a sense of trusted leadership among the leaders' followers (Park et al., 2016). While tweets allow leaders to rapidly and concisely share information, they are bound by a character limit which restricts the amount of information leaders are able to share in a single tweet. Due to this character restriction, there may be a natural tendency for leaders to post information-dense tweets and omit complimentary words that would normally frame their message more positively or negatively. Ceron et al. (2014) argue that how leaders compose and frame their tweets can be equally important as the content within the tweets. Moreover, prominent world leaders have a large following and reach on Twitter so their posts can have a big impact in setting the overall sentiment and shaping the public's perception and discussion especially during crisis situations when social media usage is heightened (Gottlieb & Dyer, 2020). If not used appropriately, social media can be detrimental to public health efforts during a public health crisis like COVID-19. In this regard, the results suggest that the G7 leaders used their ability to reach a wide audience on Twitter effectively; for instance, by being more positive towards topics such as vaccination and rebuilding future. This may be a strategic choice to influence greater participation in the vaccination programs and influence public action in accordance with rebuilding the future. Accordingly, leaders and health communicators must not only focus on the content of their posts but also actively ensure that the delivery of their tweets conveys their message as accurately as possible in a tone that is favorable during pandemics. The results also showed that the prevalence of the identified COVID-19 related topics varied over time. This is aligned with previous work; for instance, Su et al., (2021) explored topics related to COVID-19 and found considerable changes in topic prevalence over time. This may, in part, be due to the fact that pandemics are characterized by uncertainty and constantly changing scientific evidence. It follows that leaders and public health communicators should acknowledge this uncertainty and be cautious about the future merit of their messaging.

Strengths and Contributions. This work is the first to analyze the Twitter usage of G7 leaders during COVID-19 using sentiment analysis and unsupervised topic modeling. Several strengths can be attributed to this study in light of previous work. First, numerous studies have explored the public sentiment on Twitter regarding COVID-19 whereas this work explored the Twitter usage of G7 world leaders in response to COVID-19. Second, this work used a much wider selection of words to filter tweets related to COVID-19 (See Appendix A) as opposed to the majority of related studies which limit their criteria to a few keywords; for instance, "COVID-19" and "coronavirus". Third, the analysis was not limited to English language tweets and included all non-English tweets of the leaders, increasing the generalizability of the findings. Fourth, this study considered a relatively long duration of 16 months from April 2020 to July 2021 for the retrieval of tweets. Fifth, sentiment analysis was applied using multiple sentiment classification approaches including not only the VADER method which is widely used in the literature but also the Plutchik's and POMS emotion classifications. Finally, a unique feature of this analysis was the exploratory use of multivariate statistical techniques such as UMAP, t-SNE, and network analysis to glean insights from the Twitter data. In addition, this work also demonstrated the use of SHAP values to explain the proposed NLP-based multi-class topic classification model.

Limitations and Further Research. Three main recommendations for future research are proposed based on the limitations of this study. First, while the inclusion of non-English tweets is a strength of this work compared to related studies, a resulting limitation is the possible loss of semantic context during the translation of non-English tweets into English prior to analysis. Future studies may consider using multilingual sentiment libraries that directly derive sentiment based on native dictionaries of the non-English languages. Second, this work was exclusively limited to the Twitter usage of G7 world leaders. Using the methods applied in this study, future research can investigate the Twitter usage of larger groups of global leaders such as the G20 or global organizations such as the World Health Organization (WHO) or United Nations (UN). Considering a larger sample of countries or political leaders would enable deeper analysis of how country-specific socio-economic and developmental factors influence leaders' sentiments and communication styles on Twitter. Third, this study only considered data from Twitter. World leaders use several channels for crisis messaging and some leaders may have individual preferences concerning communication media. For instance, Angela Merkel of Germany was not considered in the analysis despite being part of the G7 as she did not have an active Twitter account. As such, it is acknowledged that the Twitter usage of leaders may not be fully representative of their overall crisis communication style, sentiments, and topic frames regarding COVID-19. Future research can explore the aggregation of multiple communication channels beyond Twitter for a more holistic sentiment analysis of the leaders' discussion and communication regarding COVID-19.

Concluding Thoughts. The unprecedented emergence of the COVID-19 pandemic serves as a reminder for nations, organizations, and political leaders to be prepared for inevitable future pandemics (Baum & Bangham, 2021). In this regard, platforms such as Twitter will continue to play a key role for world leaders to rapidly communicate with the public. Drawing from the sentiments, topics, and frames that emerged from the Twitter activity of the G7 leaders during COVID-19, leaders and public health communicators can better frame crisis messaging and implement communication strategies during pandemics. Ultimately, the degree to which political leaders can effectively utilize Twitter or similar trending platforms for the purposes of disseminating information and reassuring the public during pandemics would be determined not only by the content within their posts but also by the delivery and framing used to convey their messages.

Acknowledgment

I would like to thank and acknowledge:

- My supervisors - Dr. Teo Susnjak and Professor Stephen Croucher - for their technical guidance, research direction, and insightful comments and suggestions
- Massey University and PSB Academy for educational resource use
- My colleagues at Neuro - Ryan Chen, Kent Yoshimura, and the entire team - for their support and understanding
- My family, close friends, and classmates for their moral support and encouragement

Conflict of Interest Declaration

No potential conflicts of interest declared.

REFERENCES

Abd-Alrazaq, A., Alhuwail, D., Househ, M., Hai, M., & Shah, Z. (2020). Top concerns of tweeters during the COVID-19 pandemic: A surveillance study. *Journal of Medical Internet Research*, *22*(4), 1–9. https://doi.org/10.2196/19016

Ahmed, M. E., Rabin, M. R. I., & Chowdhury, F. N. (2020). COVID-19: Social Media Sentiment Analysis on Reopening. *University of Houston*. Retrieved from http://arxiv.org/abs/2006.00804

Alamoodi, A. H., Zaidan, B. B., Zaidan, A. A., Albahri, O. S., Mohammed, K. I., Malik, R. Q., … Alaa, M. (2021). Sentiment analysis and its applications in fighting COVID-19 and infectious diseases: A systematic review. *Expert Systems with Applications*, *167*, 114155. https://doi.org/10.1016/j.eswa.2020.114155

Alhajji, M., Khalifah, A., Aljubran, M., & Alkhalifah, M. (2017). Sentiment analysis of tweets in Saudi Arabia regarding governmental preventive measures to contain COVID-19. *Canopy*, *17*(2), 3–10. https://doi.org/10.20944/preprints202004.0031.v1

Barkur, G., Vibha, & Kamath, G. B. (2020). Sentiment analysis of nationwide lockdown due to COVID 19 outbreak: Evidence from India. *Asian Journal of Psychiatry*, *51*(April), 102089. https://doi.org/10.1016/j.ajp.2020.102089

Baum, J., & Bangham, C. (2021). Why disciplines must work together to prepare for future pandemics. Retrieved from: https://www.weforum.org/agenda/2021/07/disciplines-together-future-pandemics/

Behl, S., Rao, A., Aggarwal, S., Chadha, S., & Pannu, H. S. (2021). Twitter for disaster relief through sentiment analysis for COVID-19 and natural hazard crises. *International Journal of Disaster Risk Reduction*, *55*(February), 102101. https://doi.org/10.1016/j.ijdrr.2021.102101

Bernhardt, JM. (2004). Communication at the core of effective public health. Am J PublicHealth, 94(12):2051–3

Bhat, M., Qadri, M., Kundroo, M., Beg, N.-A., Ahanger, N., & Agarwal, B. (2020). Sentiment analysis of social media response on the Covid19 outbreak. *Brain, Behavior, and Immunity*, (January), 19–21.

Bird, S., Klein, E., & Loper, E. (2009). *Natural Language Processing With Python*. O'Reilly Media, Inc.

Boon-Itt, S., & Skunkan, Y. (2020). Public perception of the COVID-19 pandemic on twitter: Sentiment analysis and topic modeling study. *JMIR Public Health and Surveillance*, *6*(4), 1–17. https://doi.org/10.2196/21978

Castillo, C., Mendoza, M., & Poblete, B. (2011). Information credibility on Twitter. *Proceedings of the 20th International Conference Companion on World Wide Web, WWW 2011*, 675–684. https://doi.org/10.1145/1963405.1963500

Ceron, A., Curini, L., Iacus, S. M., & Porro, G. (2014). Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France. New Media & Society, 16(2), 340-358. doi:10.1177/1461444813480466

Colneric, N., & Demsar, J. (2020). Emotion Recognition on Twitter: Comparative Study and Training a Unison Model. IEEE Transactions on Affective Computing, 11(3), 433–446. https://doi.org/10.1109/TAFFC.2018.2807817

Credit Suisse (2018). Global Wealth Databook 2018. Credit Suisse AG Research Institute. Retrieved from: https://www.credit-suisse.com/media/assets/corporate/docs/about-us/research/publications/global-wealth-databook-2018.pdf

France Diplomacy (2018). The G7: Frequently Asked Questions. Retrieved from: https://www.diplomatie.gouv.fr/en/french-foreign-policy/french-g7-presidency-2019/the-g7-frequently-asked-questions/

Fellbaum, C. (2005). WordNet and wordnets. In: Brown, Keith et al. (eds.), Encyclopedia of Language and Linguistics, Second Edition, Oxford: Elsevier, 665-670.

Garcia, K., & Berton, L. (2021). Topic detection and sentiment analysis in Twitter content related to COVID-19 from Brazil and the USA. *Applied Soft Computing*, *101*, 107057. https://doi.org/10.1016/j.asoc.2020.107057

Ghosh, I. (2021). Visualized: Key Events in the COVID-19 Timeline. Retrieved from: https://www.visualcapitalist.com/covid-19-timeline-key-events-visualized/

Gottlieb, M., & Dyer, S. (2020). Information and Disinformation: Social Media in the COVID-19 Crisis. *Academic Emergency Medicine*, *27*(7), 640–641. https://doi.org/10.1111/acem.14036

H. Manguri, K., N. Ramadhan, R., & R. Mohammed Amin, P. (2020). Twitter Sentiment Analysis on Worldwide COVID-19 Outbreaks. *Kurdistan Journal of Applied Research*, 54–65. https://doi.org/10.24017/covid.8

Ha, H., Han, H., Mun, S., Bae, S., Lee, J., & Lee, K. (2019). An improved study of multilevel semantic network visualization for analyzing sentiment word of movie review data. Applied Sciences (Switzerland), 9(12). https://doi.org/10.3390/app9122419

Haman, M. (2020). The use of Twitter by state leaders and its impact on the public during the COVID-19 pandemic. *Heliyon*, *6*(11), e05540. https://doi.org/10.1016/j.heliyon.2020.e05540

Hutto, C.J. & Gilbert, 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

Jang, H., Rempel, E., Roth, D., Carenini, G., & Janjua, N. Z. (2021). Tracking COVID-19 discourse on twitter in north america: Infodemiology study using topic modeling and aspect-based sentiment analysis. *Journal of Medical Internet Research*, *23*(2). https://doi.org/10.2196/25431

Joseph, A., & Garside, C. (2010). The group of seven. *Canadian Packaging*, *63*(2), 22–23. https://doi.org/10.4324/9781135000356-rem203-1

Kantis, C., & Kiernan, S., Bardi, J. (2021). UPDATED: Timeline of the Coronavirus. Retrieved from: https://www.thinkglobalhealth.org/article/updated-timeline-coronavirus

Kausar, M. A., Soosaimanickam, A., & Nasar, M. (2021). Public Sentiment Analysis on Twitter Data during COVID-19 Outbreak. *International Journal of Advanced Computer Science and Applications*, *12*(2), 415–422. https://doi.org/10.14569/IJACSA.2021.0120252

Kruspe, A., Häberle, M., Kuhn, I., & Zhu, X. X. (2020). Cross-language sentiment analysis of European Twitter messages during the COVID-19 pandemic. *German Aerospace Center*. Retrieved from http://arxiv.org/abs/2008.12172

Kumar, A., Khan, S. U., & Kalra, A. (2020). COVID-19 pandemic: A sentiment analysis. *European Heart Journal*, *41*(39), 3782–3783. https://doi.org/10.1093/eurheartj/ehaa597

Lamsal, R. (2021). Design and analysis of a large-scale COVID-19 tweets dataset. *Applied Intelligence*, *51*(5), 2790–2804. https://doi.org/10.1007/s10489-020-02029-z

Larionova, M., & Kirton, J. (2020). Global Governance After the COVID-19 Crisis. *International Organisations Research Journal*, *15*(2), 7–23. https://doi.org/10.17323/1996-7845-2020-02-01

Lee, S. A. (2020). Coronavirus anxiety scale: A brief mental health screener for COVID-19 related anxiety. Death Studies, 1–9. doi:10.1080/07481187.2020.1748481

Liu, B. (2012). *Sentiment Analysis and Opinion Mining Opinion Mining. Sentiment Analysis*.

Liu, Z., Li, M., Liu, Y., & Ponraj, M. (2011). Performance evaluation of Latent Dirichlet Allocation in text mining. 2011 Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD). doi:10.1109/fskd.2011.6020066

Lu, Y., & Zheng, Q. (2021). Twitter public sentiment dynamics on cruise tourism during the COVID-19 pandemic. *Current Issues in Tourism*, *24*(7), 892–898. https://doi.org/10.1080/13683500.2020.1843607

Lundberg, & Lee, S. (2017). A Unified Approach to Interpreting Model Predictions. 31st Conference on Neural Information Processing Systems

Mathieu, E., Ritchie, H., Ortiz-Ospina, E. et al. (2021) A global database of COVID-19 vaccinations. Nat Hum Behav. https://doi.org/10.1038/s41562-021-01122-8

Maxmen, A. & Subbaraman, N. (2021). Biden's ambitious COVID plan: what scientists think. Nature 590, 18-19 (2021)

Neilson, S., Woodward, A. (2020). A comprehensive timeline of the coronavirus pandemic at 1 year, from China's first case to the present. Retrieved from: https://www.businessinsider.com/coronavirus-pandemic-timeline-history-major-events-2020-3#november-17-2019-a-person-displayed-the-first-detectable-case-of-covid-19-in-china-1

Norcross, J., Guadagnoli, E., & Prochaska, J. (1984). Factor structure of the Profile Of Mood States (POMS): Two Partial Replications. Journal of Clinical Psychology, vol. 40, no. 5, pp. 1270–1277

Park, M. J., Kang, D., Rho, J. J., & Lee, D. H. (2016). Policy Role of Social Media in Developing Public Trust: Twitter communication with government leaders. *Public Management Review*, *18*(9), 1265–1288. https://doi.org/10.1080/14719037.2015.1066418

Pasteur, L., & Koch, R. (2020). Mining Twitter Data on COVID-19 for Sentiment analysis and frequent patterns Discovery Habiba. *Algiers University*, *74*(1934), 535–546.

Plutchik, R. (1980) A General Psychoevolutionary Theory of Emotion. Theories of Emotion. Academic Press, vol. 1, pp. 3–33.

Pokharel, B. P. (2020). Twitter Sentiment Analysis During Covid-19 Outbreak in Nepal. *SSRN Electronic Journal*, (March), 1–9. https://doi.org/10.2139/ssrn.3624719

Puri, N., Coomes, E. A., Haghbayan, H., & Gunaratne, K. (2020). Social media and vaccine hesitancy: new updates for the era of COVID-19 and globalized infectious diseases. *Human Vaccines and Immunotherapeutics*, *16*(11), 2586–2593. https://doi.org/10.1080/21645515.2020.1780846

Rahman, M. M., Ali, G. G. M. N., Li, X. J., Samuel, J., Paul, K. C., Chong, P. H. J., & Yakubov, M. (2021). Socioeconomic factors analysis for COVID-19 US reopening sentiment with Twitter and census data. *Heliyon*, *7*(2), e06200. https://doi.org/10.1016/j.heliyon.2021.e06200

Ratzan, S. C., Sommariva, S., & Rauh, L. (2020). Enhancing global health communication during a crisis: Lessons from the COVID-19 pandemic. *Public Health Research and Practice*, *30*(2). https://doi.org/10.17061/phrp3022010

Rufai, S. R., & Bunce, C. (2020). World leaders' usage of twitter in response to the COVID-19 pandemic: A content analysis. *Journal of Public Health (United Kingdom)*, *42*(3), 510–516. https://doi.org/10.1093/pubmed/fdaa049

Samuel, J., Rahman, M. M., Ali, G. G. M. N., Samuel, Y., Pelaez, A., Chong, P. H. J., & Yakubov, M. (2020). Feeling Positive about Reopening? New Normal Scenarios from COVID-19 US Reopen Sentiment Analytics. *IEEE Access*, *8*, 142173–142190. https://doi.org/10.1109/ACCESS.2020.3013933

Sanders, A. C., White, R. C., Severson, L. S., Ma, R., McQueen, R., Paulo, H. C. A., … Bennett, K. P. (2020). Unmasking the conversation on masks: Natural language processing for topical sentiment analysis of COVID-19 Twitter discourse. *MedRxiv*, 2020.08.28.20183863. Retrieved from https://doi.org/10.1101/2020.08.28.20183863

Sengupta, S., Mugde, S., & Sharma, G. (2020). An exploration of impact of COVID 19 on mental health - Analysis of tweets using natural language processing techniques. *MedRxiv*, 436286.

Chen, S. (2021). Interpretation of multi-label classification models using shapley values. Mathematics for Engineers, Institue for Technologies of Metals, University of Duisburg-Essen

Singh, P., Singh, S., Sohal, M., Dwivedi, Y., Singh, K., & Sawhney, R. (2020). Psychological fear and anxiety caused by COVID-19: Insights from Twitter analytics. *Asian Journal of Psychiatry*, (January).

Solangi, Y. A., Solangi, Z. A., Aarain, S., Abro, A., Mallah, G. A., & Shah, A. (2019). Review on Natural Language Processing (NLP) and Its Toolkits for Opinion Mining and Sentiment Analysis. *2018 IEEE 5th International Conference on Engineering Technologies and Applied Sciences, ICETAS 2018*, 1–4. https://doi.org/10.1109/ICETAS.2018.8629198

Statista (2021). Number of monetizable daily active Twitter users (mDAU) worldwide from 1st quarter 2017 to 2nd quarter 2021. Retrieved from: https://www.statista.com/statistics/970920/monetizable-daily-active-twitter-users-worldwide/

Su, Y., Venkat, A., Yadav, Y., Puglisi, L. B., & Fodeh, S. J. (2021). Twitter-based .analysis reveals differential COVID-19 concerns across areas with socioeconomic disparities. *Computers in Biology and Medicine*, *132*(March), 104336. https://doi.org/10.1016/j.compbiomed.2021.104336

Villavicencio, C., Macrohon, J. J., Inbaraj, X. A., Jeng, J.-H., & Hsieh, J.-G. (2021). Twitter Sentiment Analysis towards COVID-19 Vaccines in the Philippines Using Naïve Bayes. *Information*, *12*(5), 204. https://doi.org/10.3390/info12050204

Wild, K. (2020). Using corpora to track the language of Covid-19: update 2. Oxford English Dictionary (OED). Retrieved from: https://public.oed.com/blog/using-corpora-to-track-the-language-of-covid-19-update-2/

Xue, J., Chen, J., Hu, R., Chen, C., Zheng, C., Su, Y., & Zhu, T. (2020). Twitter discussions and emotions about the COVID-19 pandemic: Machine learning approach. *Journal of Medical Internet Research*, *22*(11), 1–14. https://doi.org/10.2196/20550

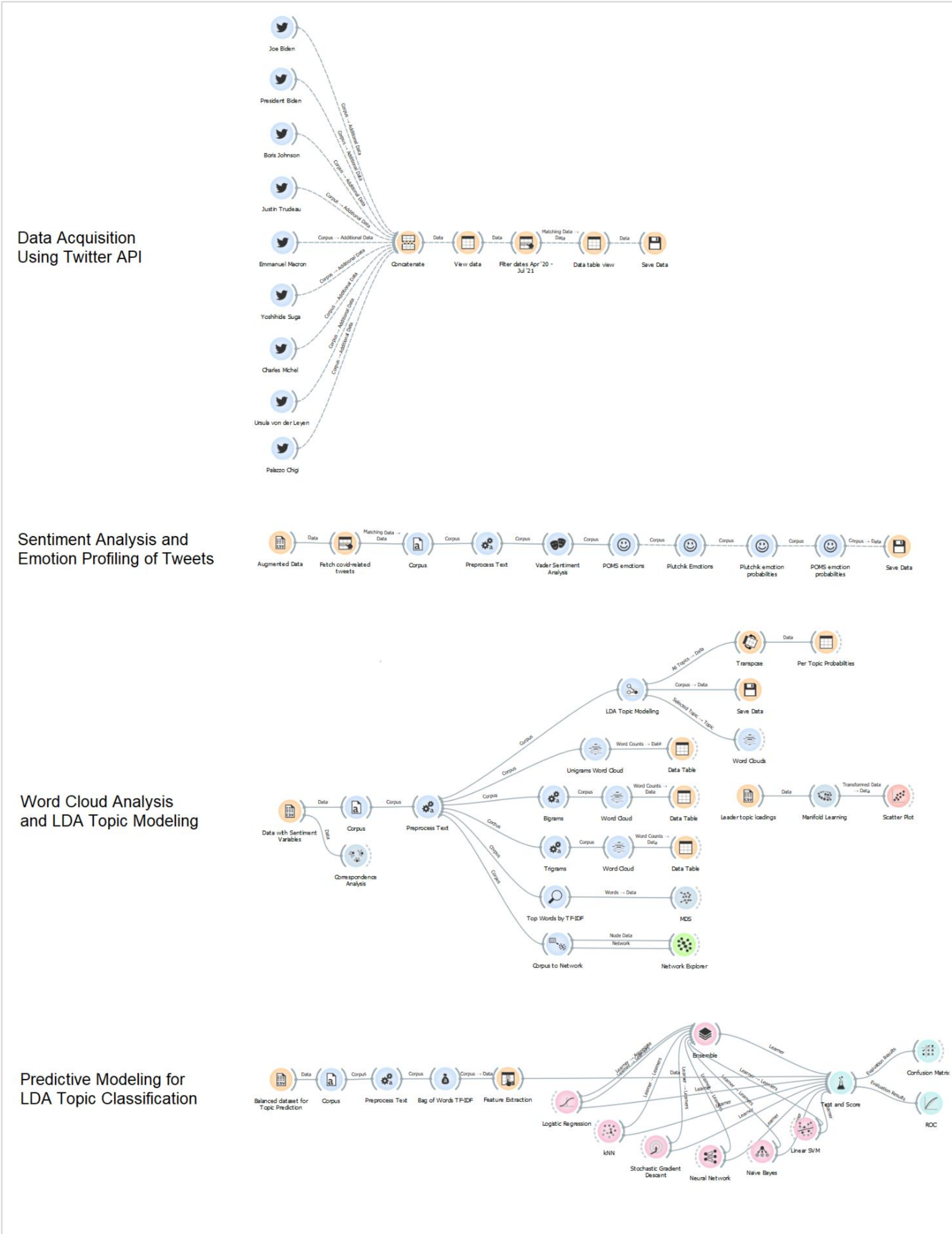
Ye, N. (2013). The Handbook of Data Mining. Routledge.

APPENDICES

Appendix A – COVID-19 Related List of Words Used to Filter Tweets (Wild, 2020)

covid	*corona*	*supply*
covid-19	*infect*	*recover*
covid19	*test*	*support*
coronavirus	*swab*	*resilien*
vaccin	*diagnos*	*situation*
mask	*illness*	*track*
pandemic	*mild*	*died*
stay safe	*reopen*	*crisis*
lockdown	*essential*	*back*
quarantine	*frontline*	*dose*
epidemic	*face*	*Pfizer*
vaccine	*isolat*	*Astra*
vaccination	*health*	*Moderna*
death	*hospital*	*mRNA*
cases	*distance*	*Sinovac*
SARS	*distancing*	*flu*
infection	*PPE*	*protect*
contagious	*safe*	*aid*
spread	*meter*	*help*
virus	*antibody*	*inoculat*
outbreak	*ventilator*	*pharma*
pneumonia	*oxygen*	*doctor*
nCov	*N95*	*nurse*
Wuhan	*ICU*	*ban*
disease	*shots*	*gradual*
symptom	*economy*	*open*
isolation	*bounce*	*fight*
immun	*measure*	*battle*
curve	*strong*	*together*
social distanc	*hope*	*transmis*
delta	*survive*	*case*
variant	*medical*	*strength*
community	*booster*	*remote*

Appendix B – Analysis Workflows



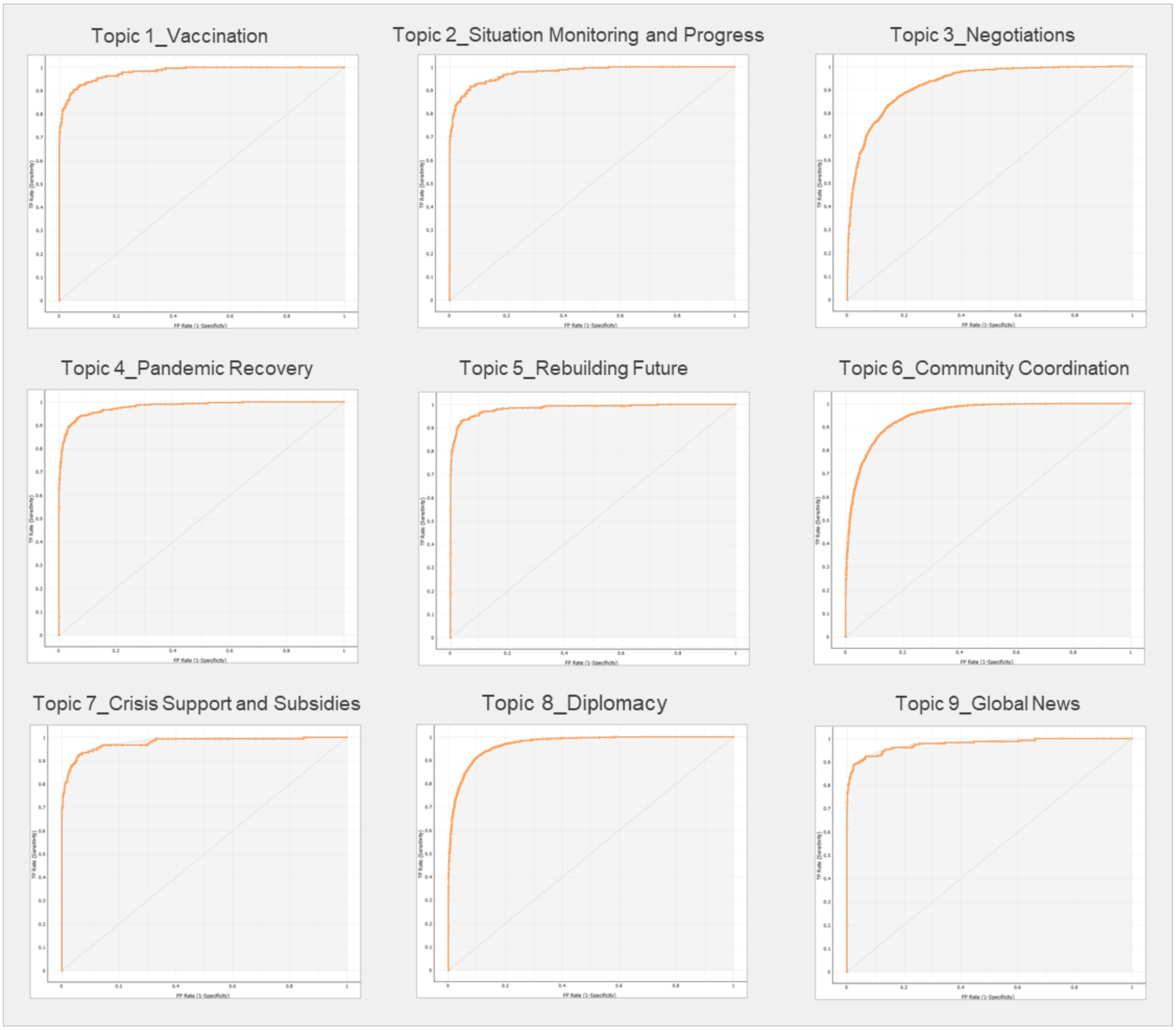
Appendix C – Variables in Final Dataset (N=9918)

Variable Theme	Variable	Type	Description
Twitter Data	*date*	datetime	Tweet date (Apr 01, 2020 - Jul 31 2021)
	month	datetime	Month of tweet
	year	datetime	Year of tweet
	day	datetime	Day of tweet
	month_year	datetime	Month and Year of tweet
	weekday	nominal	Weekday of tweet
	username	nominal	Author of tweet
	leader	nominal	G7 leader name
	tweet	text	Tweet text
	translated_tweet	text	Translated tweet text
	country	nominal	Country of G7 leader
	language	nominal	Language of tweet
	length	continuous	Number of characters in tweet
	like_count	continuous	Number of tweet likes
	retweet_count	continuous	Number of retweets
	author_status_count	continuous	Author status count as of date
	author_favorites_count	continuous	Author favorities count as of date
	author_following_count	continuous	Author following count as of date
	author_follower_count	continuous	Author follower count as of date
G7 Country-level variables	*cty_new_cases_per_mil*	continuous	New COVID-19 cases per million
	cty_reproduction_rate	continuous	COVID-19 virus reproduction rate
	cty_new_vaccinations	continuous	Number of new COVID-19 vaccinations
	cty_fully_vaccinated_per_hundred	continuous	Fully vaccinated per 100 as of date
	cty_stringency_index	continuous	COVID-19 Stringency index
	cty_covid_related_anxiety	continuous	Average COVID-19 anxiety level
	cty_population_density	continuous	Population density
	cty_gdp_per_capita	continuous	GDP per capita
	cty_hosp_beds_per_thousand	continuous	Hospital beds available per thousand
	cty_human_development_index	continuous	Human development Index
	cty_social_support_index	continuous	Social support index
	cty_corruption_index	continuous	Corruption index
	cty_happiness_index	continuous	Happiness Index
	cty_GINI_index	continuous	GINI index for income inequality
Sentiment Analysis	*positive*	continuous	Level of positive sentiment in tweet
	negative	continuous	Level of negative sentiment in tweet
	neutral	continuous	Level of neutral sentiment in tweet
	polarity	continuous	Sentiment polarity of tweet
	subjectivity	continuous	Subjectivity of tweet
	vader_sentiment	nominal	Positive, negative, or neutral classification
	POMS_emotion_class	nominal	POMS emotion profile of tweet
	Plutchik_emotion_class	nominal	Plutchik emotion profile of tweet
	Plutchik_Anger	continuous	Weight for Plutchik_Anger
	Plutchik_Disgust	continuous	Weight for Plutchik_Disgust
	Plutchik_Fear	continuous	Weight for Plutchik_Fear
	Plutchik_Joy	continuous	Weight for Plutchik_Joy
	Plutchik_Sadness	continuous	Weight for Plutchik_Sadness
	Plutchik_Surprise	continuous	Weight for Plutchik_Surprise
	Plutchik_Trust	continuous	Weight for Plutchik_Trust
	Plutchik_Anticipation	continuous	Weight for Plutchik_Anticipation
	POMS_Anger	continuous	Weight for POMS_Anger
	POMS_Depression	continuous	Weight for POMS_Depression
	POMS_Fatigue	continuous	Weight for POMS_Fatigue
	POMS_Vigour	continuous	Weight for POMS_Vigour
	POMS_Tension	continuous	Weight for POMS_Tension
	POMS_Confusion	continuous	Weight for POMS_Confusion
Dummy Variables for Keywords	*contains_mask*	binary	1 if tweet contains "mask", else 0
	contains_vaccin	binary	1 if tweet contains "vaccin", else 0
	contains_travel	binary	1 if tweet contains "travel", else 0
	contains_quarantine	binary	1 if tweet contains "quarantine", else 0
	contains_reopen	binary	1 if tweet contains "reopen", else 0
	contains_econom	binary	1 if tweet contains "econom", else 0
	contains_delta	binary	1 if tweet contains "delta", else 0
	contains_future	binary	1 if tweet contains "future", else 0
LDA Topic Modeling	*Topic 1*	continuous	Probability for LDA Topic 1
	Topic 2	continuous	Probability for LDA Topic 2
	Topic 3	continuous	Probability for LDA Topic 3
	Topic 4	continuous	Probability for LDA Topic 4
	Topic 5	continuous	Probability for LDA Topic 5
	Topic 6	continuous	Probability for LDA Topic 6
	Topic 7	continuous	Probability for LDA Topic 7
	Topic 8	continuous	Probability for LDA Topic 8
	Topic 9	continuous	Probability for LDA Topic 9
	Dominant_Topic	nominal	LDA Topic with highest probability

Appendix D – Descriptive Statistics of Country Indicators by G7 Country (Apr'20 – Jul'21)

Indicator	Canada	EU	France	Italy	Japan	United Kingdom	United States	Overall
mean_new_cases_per_mil	109.90	192.77	184.01	139.28	19.22	161.32	195.87	160.14
mean_reproduction_rate	0.92	1.02	1.07	0.99	1.09	1.02	1.03	1.01
mean_new_vaccinations	264486	1135150	177033	165998	315574	175186	644170	520033
max_fully_vaccinated_per_hundred	58.95	46.05	46.70	51.20	29.33	56.14	49.17	58.95
mean_stringency_index	73.51	59.65	61.77	69.46	43.66	69.84	64.81	65.22
mean_covid_related_anxiety	12.76	11.62	11.24	11.25	12.43	12.41	13.21	12.33
mean_population_density	4.04	178.67	122.58	205.86	347.78	272.90	35.61	117.91
mean_gdp_per_capita	44017	37182	38605	35220	39002	39753	54225	43499
mean_hosp_beds_per_thousand	2.50	4.99	5.98	3.18	13.05	2.54	2.77	3.94
mean_human_development_index	0.93	0.90	0.90	0.89	0.92	0.93	0.93	0.92
mean_social_support_index	0.93	0.92	0.95	0.89	0.89	0.93	0.94	0.93
mean_corruption_index	0.43	0.63	0.57	0.84	0.61	0.49	0.68	0.60
mean_happiness_index	7.32	6.39	6.44	5.96	5.92	6.71	6.99	6.72
mean_GINI_index	33.30	31.85	31.60	35.90	32.90	34.80	41.10	35.30

Appendix E – ROC Curves by LDA Topic Class (Ensemble Model)



Appendix F – Non-meaningful Results of k-Means Clustering for Topic Detection (Discarded from Analysis due to lack of interpretability)