Information Systems Frontiers
https://doi.org/10.1007/s10796-024-10542-6

Navigating in Turbulent Times: Using Social Media to Examine Small and family-Owned Business Topics and Sentiments during the COVID-19 Crisis

Shaun Meric Menezes[1] · **Ashok Kumar**[1] · **Shantanu Dutta**[1]

Accepted: 13 September 2024

Abstract

During a crisis, small and family-owned businesses tend to experience more severe economic consequences than their larger counterparts and often lack financial resources needed to weather the challenges brought about by the crisis. To comprehend the distinct challenges and concerns of small and family-owned businesses during a major crisis, this research study focuses on the recent COVID-19 pandemic, which had a catastrophic effect on businesses and societies alike. To that effect, we address two research questions: First, what topics pertaining to small and family-owned businesses do social media users discuss during the COVID-19 pandemic? To achieve this goal, we employ the BERTopic model, a state-of-the-art technique for topic modeling, to identify and categorize prevalent themes arising from the discourse. Second, what is the impact of major government announcements on these discussions? Specifically, we study how sentiments change around a major government announcement aimed at supporting small businesses in the face of the pandemic. Our findings suggest that government announcements do not change the negative sentiments for most of the topics. This highlights the ineffectiveness of government announcements in alleviating people's concern related to small and family-owned business and underscores the importance of a better consultation process and communication strategy by policymakers. The implications of our study transcend recent COVID-19 effects, as World Health Organization (WHO) cautions that there could be even worse health and socio-economic crises in the future, and we need to be better prepared to handle subsequent devastating effects.

Keywords Coronavirus · COVID-19 · Topic model · BERTopic · Clustering · Social media

1 Introduction

During various crises, small and family-owned businesses often encounter significant challenges due to their inherent vulnerabilities such as inadequate planning, restricted cash flow, limited access to capital, lack of government assistance, and infrastructure deficiencies (Runyan, 2006). A few studies have shown that small businesses face disproportionately higher level of challenges relative to larger firms amid crisis periods (see for example, Sahin et al. (2011) on 2007–2009 financial crisis; Marshall et al. (2015) on 2005 Hurricane Katrina; Wedawatta et al. (2013) on 2019 Cockermouth flood; McNamara (2013) on 2012 flood in Fiji; Irvine and Anderson (2006) on 2001 foot and mouth disease in Scotland and England). Even in the context of manmade crises, such as the 2011 London riots, research by Doern (2014) revealed that affected businesses faced substantial financial losses, alongside employee layoffs, reduced investments, and enduring personal and psychological costs.

Transitioning from this historical backdrop, the COVID-19 pandemic has presented unprecedented challenges across various domains, primarily posing a significant threat to the lives of millions globally (Dutta et al., 2023; Worldometers, 2023; Belitski et al., 2021), and significantly impacting

✉ Shantanu Dutta
shantanu.dutta@telfer.uottawa.ca

Shaun Meric Menezes
smene050@uottawa.ca

Ashok Kumar
akuma127@uottawa.ca

[1] Telfer School of Management, University of Ottawa, 55 Laurier Ave E, Ottawa, ON K1N 6N5, Canada

Published online: 18 October 2024

economies and individuals of countries (Fairlie & Fossen, 2021; McKibbin & Fernando, 2020; Hadjielias et al., 2022). Nations worldwide swiftly implemented a range of measures, including lockdowns, quarantines, curfews, and mobility restrictions, in a pressing effort to contain the virus (Brülhart et al., 2020; Onyeaka et al., 2021).

While smaller and family-owned businesses employ more than 60% of the US workforce and contribute approximately 60% of US GDP (FamilyBusiness, 2021, June 2), we often do not pay adequate and systematic attention to understand their challenges and incorporate their concerns in policy decisions. For example, the impact of COVID-19 pandemic has been uneven, with self-employed individuals bearing a disproportionate burden on their income compared to salaried employees (Kritikos et al., 2020; Hadjielias et al., 2022) and certain employees, particularly those working for small businesses, faced layoffs as these businesses grappled with insufficient liquidity to maintain their workforce amidst social distancing and lockdowns (Kalogiannidis, 2020). The distinct challenges encountered by smaller and family-owned business during the pandemic thus come to the fore (Tang et al., 2021). This study attempts to understand the sentiments of people related to small and family businesses during pandemic, and how they reacted to the tailored policy interventions to address their specific needs.

During COVID-19 pandemic, when most of the people were confined to their homes, social media (such as Twitter[1], Facebook, Reddit) became a useful source to gauge their sentiment (Cho et al., 2023; González-Padilla & Tortolero-Blanco, 2020; Lips, 2021; Mendon et al., 2021; Choudrie et al., 2021). Accordingly, we rely on social media (specifically, Twitter) to explore the topics discussed by the users in context of small and family-owned businesses and how various government announcements impacted the sentiment of the users. To that effect, using Twitter data, we explore the following two research questions: First, what topics pertaining to small and family-owned businesses do social media users discuss during the COVID-19 pandemic? The primary objective of this inquiry is to discern the emerging topics within discussions concerning small and family business operations amid the COVID-19 pandemic on Twitter. To achieve this goal, we employ the BERTopic model, a state-of-the-art technique for topic modeling (Egger & Yu, 2022; Grootendorst, 2022; Tang et al., 2024; Williams et al., 2023; Ruocco et al., 2024), to identify and categorize prevalent themes arising from the discourse. Second, what is the impact of major government announcements on these discussions? Specifically, we study how sentiments change around a major government announcement aimed at supporting small businesses in the face of the pandemic. This will highlight the effectiveness of relevant government interventions in alleviating negative concerns of affected people.

In summary, our research aims to uncover emerging topics within small and family-owned business discussions during COVID-19 on Twitter using the BERTopic model and to track the changes in sentiments of the twitter users around the major government announcements. This investigation is important and makes significant contributions at least for three reasons. First, this study contributes to the growing body of knowledge concerning the issues and difficulties faced by small and family-owned businesses during crises, such as the recent conclusion of the COVID-19 pandemic. Earlier studies did not pay much attention to small and family-owned firms in the context of COVID-19. Second, we examine the role and influence of government announcements and policies in shaping social media discussions and mitigating the challenges faced by small and family-owned businesses during this global crisis. Third, our study contributes to the growing body of literature that focuses on social media analysis to aid business and policy decisions (Bhatt et al., 2022) by leveraging recent advancements in the field of Natural Language Processing (NLP). We would like to also highlight that implications of our study transcend recent COVID-19 effects. Although the World Health Organization (WHO) officially ended the COVID-19 global health emergency in early 2023, WHO cautions that future health crises could be even worse. According to WHO, there is a growing concern surrounding the potential emergence of another severe pandemic, referred to as "Disease X" which could have devastating affects around the globe (WHO, 2023, May 22; Reuters, 2023, August 18). Therefore, relevance of our findings goes beyond the scope of recently concluded pandemic.

In Sect. 2, we review the relevant literature on the potential challenges faced by smaller and family-owned businesses during the COVID-19 pandemic. This is followed by Sect. 3, where we present our methodology, which includes the sample and the topic modeling technique that is most suitable for our analysis and rationale behind it. Section 4 presents the results of our analysis. In Sect. 5, we discuss the summary of research findings, and in Sect. 6, we discuss the policymaking and practical implications of our study. Finally, Sect. 7 concludes the study.

2 Relevant Literature and Theoretical Background

In this section, firstly, under the theme of 'small and family-owned business literature review' (Sect. 2.1), we explore four interconnected streams of literature related to small

[1] Twitter now has been rebranded as X.

and family-owned businesses, establishing the theoretical foundation of our study. The initial segment explores the general challenges that small and family-owned businesses encountered, coupled with the psychological burdens borne by their proprietors. Subsequently, the second part delves into the specific impact on women-owned small businesses. The third part discusses the influence of government interventions in assisting small and family-owned businesses. Finally, the fourth part examines the issue of adaptability in the face of a new working culture within small and family-owned businesses. These segments would help us to contextualize the findings of the topic modelling, as discussed later. Thereafter, we summarize the literature on the textual analysis of social media discussion (Sect. 2.2) and the usefulness of the 'Twitter' as an important social media platform (Sect. 2.3) and conclude by discussing theoretical background (Sect. 2.4).

2.1 Small and Family-Owned Business Literature Review

The repercussions of COVID-19 distinctly affected small businesses (Bruhn et al., 2023), resulting in an unprecedented surge in the demand for loans (Cowling et al., 2021; Bartik et al., (2020). However, only a few studies have examined the effect of COVID-19 pandemic on smaller and family-owned businesses. For instance, Alekseev et al. (2023) conducted a survey in 2000 across small business owners, managers, and employees to analyze the repercussions of the COVID-19 pandemic on these enterprises. They find that young small business struggled to keep their business open during the pandemic, such firms had to lower the prices to address demand shock and did not receive adequate financial support. Using a sample of UK based small businesses over 2016–2019, Belghitar et al. (2021) examined the impact of government policies on small businesses and their subsequent responses. They show that government funding was critical for the survival of most of the smaller businesses. Stephens et al. (2021) conducted a survey in on a small sample ($n = 113$) to understand the impact of COVID-19 on female entrepreneurs. In the context of our study, we would like to explore how these issues are manifested in the topic modelling analysis.

2.1.1 Psychological Impact and Workforce Challenges

During the COVID-19 period, Patel and Rietveld (2020) observed that self-employed individuals reported heightened psychological distress due to financial insecurity. Concurrently, Chhatwani et al. (2022) suggested that depression may impede the beneficial influence of psychological resilience on a business's ability to persevere and thrive in the face of adversities. Hence, it may be evident that these mental obstacles may potentially lead to detrimental impacts on the overall survival and operational efficiency of these businesses and might be important to understand the sentiments before and after the announcements.

2.1.2 Challenges Confronting Women-Owned Small Businesses

Prior to the pandemic, Ackah et al. (2023) conducted a study in Ghana, which revealed that women-owned businesses faced a productivity gap in comparison to businesses owned by men. However, the introduction of the pandemic had significant effects on small businesses owned by women (Manolova et al., 2020). In the UK, Dy and Jayawarna (2020) highlighted how the pandemic disproportionately impacted self-employed women and women-owned businesses, particularly those from marginalized backgrounds. Furthermore, research conducted in Ireland Stephens et al. (2021) revealed that certain female entrepreneurs demonstrated resilience by leveraging peer support. These studies collectively underline the significance of delving into the genuine challenges encountered by women-owned small businesses during the pandemic and understanding these dynamics may be essential for addressing the specific challenges women-owned small and family businesses face during crises, providing a foundation for targeted social media and policy interventions.

2.1.3 Government Support: Navigating Pandemic Challenges for Small and Family-Owned Businesses

As the pandemic led to lockdowns and restrictions, certain financially fragile small businesses initially encountered demand shocks, prompting them to lower their prices. However, as time progressed, supply shocks took precedence over the earlier concerns about demand (Alekseev et al., 2023). Governments had to intervene to provide essential support and mitigate the economic impact through various policy measures such as stimulus packages, grants, and loan programs, governments aimed to provide financial relief, ensure business continuity, and protect jobs (Belghitar et al., 2021; Dörr et al., 2021). These interventions were crucial for enabling small businesses to navigate the uncertainty and financial hardships brought about by the pandemic, ultimately contributing to the preservation of the economic fabric of communities (Dörr et al., 2021). This analysis suggests the need for a comprehensive understanding of how social media conversations may reflect and influence the resilience and recovery strategies of small and family-owned businesses during crises, which could enhance the relevance of targeted support and policy adjustments.

2.1.4 Navigating Adaptability Challenges in Small and Family-Owned Businesses amidst Change

The rapid transition to remote work due to the pandemic posed adaptability challenges for businesses, impacting their work processes and social dynamics (Lal et al., 2021). Zhang et al. (2021) and Wilton et al. (2011) also highlighted that as a result of this disruption, hybrid or work-from-home (WFH) oriented workplaces emerged, necessitating organizational adaptations. In response, family businesses proactively addressed these challenges by leveraging their existing digital infrastructure and embarking on a steep learning curve (Soluk et al., 2021). Studying these adaptive measures and the resulting shifts in work culture within small and family-owned businesses may be crucial for understanding the long-term implications of the pandemic or any upcoming catastrophes on the business landscape and my help in informing future strategies for resilience and innovation.

2.2 Textual Analysis of Social Media Discussion

Platforms like Twitter and Facebook have been widely used as sources of information to study health-related issues (Pollack et al., 2023; Asghari et al., 2018; Mejova et al., 2018; Weeg et al., 2015). These user-generated posts on social media have become a valuable data source and have caught the attention of many researchers (Lin et al., 2020; Wang et al., 2020; Yin et al., 2022; Xue et al., 2020). Studies have leveraged a combination of topic modeling and sentiment analysis to understand public opinion and discourse on social media (Corti et al., 2022; Rahmadan et al., 2020; Yin et al., 2022). Rahmadan et al. (2020) applied Latent Dirichlet Allocation (LDA) to identify topics in tweets about the Jakarta flood disaster, revealing predominantly negative sentiments focused on flood impacts and public feedback on disaster management. Similarly, Corti et al. (2022) used topic modeling to analyze tweets and understand the emotional stance about autism spectrum disorder (ASD) during the COVID-19 pandemic, highlighting the support within the ASD community and the discussion of ASD in the context of the pandemic. Employing NLP techniques, Yin et al. (2022) conducted an in-depth analysis of Twitter discussions about the COVID-19 vaccine. These studies demonstrate the utility of combining topic modeling and sentiment analysis to gain insights into public discourse on critical issues as expressed on social media platforms.

2.3 Why Twitter?

Twitter is a well-known social media platform where people express their thoughts about specific topics using short written messages called “tweets” (D’Andrea et al., 2019). A tweet also includes information about where the person is located, hashtags (keywords used to categorize content), and emoticons (symbols representing emotions), which all help to show how people are feeling about certain topics or issues (Giachanou & Crestani, 2016). Additionally, government officials use Twitter to share information about events or important announcements with the general public (Golbeck et al., 2010). Researchers have found Twitter to be very useful for various types of studies. For example, it has been used to study the opinion of flight services (Tiwari et al., 2019), the feelings people have about sports (Yu & Wang, 2015), what people think about politics (Khatua et al., 2020), the emotions of cancer patients (Crannell et al., 2016), and people’s opinions about vaccines (D’Andrea et al., 2019).

Twitter provides researchers with a special collection of text that offers advantages over traditional methods (such as surveys and interviews) of collecting information. One significant advantage is that Twitter data is updated in real-time, enabling researchers to observe current discussions and track trends as they happen. This real-time aspect is crucial for studying dynamic topics, such as the response of small and family-owned businesses to the COVID-19 pandemic and the shift of public sentiment in relation to key government announcements. Therefore, analyzing Twitter data is highly relevant to our study. It allows us to identify and explore various topics related to small and family-owned businesses during the COVID-19 pandemic and to examine how public sentiment changes around significant government announcements. This approach provides a comprehensive view of the discourse and sentiment trends, offering valuable insights into the challenges and perceptions faced by these businesses during a critical period. One of the underlying assumptions in our study is that tweets reflect social emotions of the users adequately. Accordingly, we discuss the theoretical background that highlights the usefulness of tweets in sharing emotions on socio-economic issues, in the next section.

2.4 Theoretical Background

The advancement of information technology has led to digital means of knowledge sharing and interaction, primarily through online social media platforms like Twitter (Panahi et al., 2015). This transformation has accelerated the dissemination and real-time updating of knowledge (Panahi et al., 2016). As posited by Papadopoulos et al. (2013) and Chang and Chuang (2011), social media serves as a collaborative online environment where individuals actively exchange information and knowledge (Chu et al., 2015). Previous studies have highlighted the significance of sharing emotions on Twitter and similar platforms, noting that this type of social sharing acts as a form of interpersonal interaction that assists individuals in managing their emotions (Hidalgo et al., 2017; Bachura et al., 2022).

Emotions arise from an individual's interaction with their environment and involve complex, coordinated responses (Scherer, 2005). When someone experiences an emotionally significant event, they quickly express their emotions to those around them, initiating a process known as Social Sharing of Emotion (SSE) (Rimé, 2009; Luminet et al., 2000). SSE is inherently a social activity, requiring at least two people: (1) the person experiencing the emotion, who feels compelled to share it, and (2) the recipient of this emotional expression. Research has shown that SSE occurs across all types of emotions, genders, ages, cultures, and educational levels, although there are slight variations among these groups (Rimé et al., 1998).

The Social Sharing of Emotion (SSE) theoretical framework posits that individuals communicate openly with one or more persons about the circumstances of an emotion-eliciting event and their own feelings and emotional reactions (Rimé, 2009). People often share their emotions on social media, which scientists use to assess public sentiment during crises. Research indicates that negative emotions such as fear, anger, sadness, and insecurity are common in these situations (Steinert, 2020). Sharing emotions online does not necessarily mean these feelings are inauthentic or unrepresentative of true feelings. While some individuals might misrepresent their emotions online, there is no evidence to suggest that this is widespread (Steinert, 2020). During the COVID-19 pandemic, many people felt that their valued aspects of life were under threat, leading to a prevalence of negative emotions (Steinert, 2020).

Inspired by the SSE framework, we uncover the themes of Twitter discussions related to small and family-businesses during COVID-19 and conduct statistical analyses to understand how announcement declarations are received by relevant Twitter users. Specifically, we aim to determine if government announcements can alleviate people's negative concerns, given that sharing negative emotions has been shown to be beneficial when the SSE framework is met with appropriate feedback (Rimé, 2009; Hidalgo et al., 2015). This study contributes to the extension of SSE framework in specific contexts, including crisis situations, small businesses, and entrepreneurship, thereby enhancing our understanding of the utility of social networks as platforms for emotional expression, which has become increasingly prevalent in contemporary times.

3 Methodology

In our research, we utilized topic modelling to uncover prevalent themes within Twitter discussions concerning small and family-owned businesses during the COVID-19 pandemic. The methodology is illustrated in Fig. 1. The data collection process involved gathering a dataset of 132,567 tweets, guided by specific keywords associated with small and family-owned businesses (refer to Table 1). Prior to analysis, the collected tweets underwent text pre-processing to enhance topic exploration.

We employed BERTopic, a powerful topic modeling technique (Egger & Yu, 2022; Grootendorst, 2022; Tang et al., 2024; Williams et al., 2023; Ruocco et al., 2024) that leverages deep learning models to capture complex relationships within textual content. BERTopic allows us to extract meaningful clusters and uncover underlying themes in the tweets. By comparing the outcome of BERTopic to that of other traditional topic modeling algorithms like LDA and NMF, we find that BERTTopic performs better in capturing contextual connections between words, resulting in more accurate topic identification (Röder et al., 2015; Lau et al., 2016). Online Appendix[2] 1 presents the topics generated by LDA and NMF algorithms, respectively. The application of BERTopic's hierarchical structure and adaptability to diverse data types ensured a comprehensive analysis of the topics discussed in small and family-owned businesses during the pandemic. Initially, the implementation of BERTopic yielded 18 meaningful topics, as presented in Online Appendix 2, Table A2.1. However, through the process of grouping and merging similar topics, the number of topics was reduced to 11 as shown in Table A2.3, which formed the basis for our subsequent analysis.

Additionally, we explored the effectiveness of BERTopic by analyzing its performance in handling outliers – tweets that did not align with topics identified by BERTopic. Employing cosine similarity, we aimed to reduce these outliers and assign tweets to suitable thematic topics. This approach did not drastically increase the tweet count per topic, except one topic out of 11 topics, highlighting BERTopic's adaptability and effectiveness in identifying meaningful topics from Twitter conversations and leaving out the tweets which does not pertain to these topics.

In our exploration of shifting sentiments around major government announcements, we adopted a structured framework to analyze the change of sentiments post a major government announcement during COVID-19, in relation with topics identified by BERTopic. Motivated by other studies (Lee et al., 2016; Choi & Kim, 2013; Yu et

[2] Online Appendix can be accessed using the following link: https://drive.google.com/drive/folders/17nrbgagLpAegO4vWj6iQsWUDZbjSdL6v?usp=drive_link.

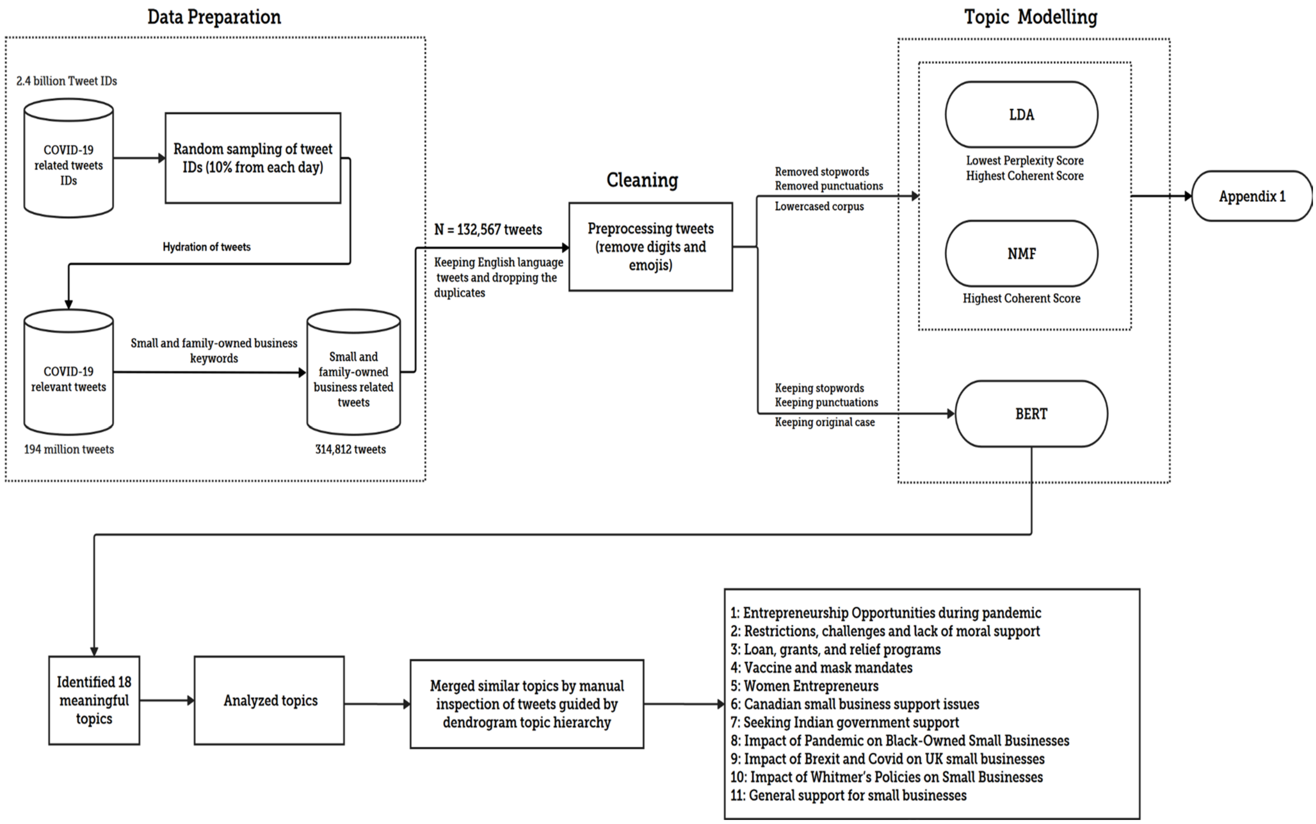

Fig. 1 Process model

Table 1 The keywords employed for retrieving tweets related to small and family-owned businesses

Small and family-owned business keywords
family business, household enterprise, kinship company, domestic firm, entrepreneur, entrepreneurs, entrepreneurship, small business, mom-and-pop shop, mom and pop shop, microenterprise, sole proprietorship, family firm, ancestral business, hereditary enterprise, progenitor company, bloodline enterprise, familial company, household venture, family-owned operation, family owned operation, family owned business, blood relation corporation, family-run business, family run business

al., 2023), we have used Linguistic Inquiry and Word Count (LIWC) software developed by Pennebaker et al. (2007) to measure the extent of negative tone (tone_neg) and anxiety levels (emo_anx) in our texts around key government announcements[3]. Later we performed t-test to examine the change in the negative tone and anxiety levels around the key announcement dates. This multi-step process involved data collection, identification of key announcement dates, temporal window selection, topic-centric filtering, and documenting the effectiveness of the announcements by using the t-test results.

[3] The reason for focusing on negative tone and emotional anxiety is that people often share the negatively connotated emotions on social media during a crisis (Steinert, 2020).

3.1 Data Description

For this study, a substantial dataset comprising over 2.4 billion tweet IDs (specifically, 2,441,061,762 tweet IDs) related to the COVID-19 topic was sourced from the echen102 GitHub repository (Chen et al., 2020). This dataset covered each day in the timeline from January 2020 to December 2022. A systematic approach of random sampling was adopted to select a representative 10% of tweets from each day. After "hydration" – a process to retrieve complete tweet data using only the tweet IDs and the Twitter API – a final dataset of 194,207,066 tweet IDs were collected. This resulted in the successful construction of a comprehensive dataset, including complete tweets, author details, engagement metrics (such as likes, retweets, tweet count, and number of followers). We presented a discussion on the metadata plots in Section A2.3 of Online Appendix 2.

Further, we utilized this extensive dataset to explore the topics revolving around small and family-owned businesses on Twitter, employing carefully selected keywords (as outlined in Table 1), we systematically extracted tweets relevant to small and family-owned businesses from the dataset. This filtration process yielded a focused subset of 314,812 tweets. After considering only English language tweets and eliminating duplicates, we end up with a total of

132,567 tweets as our final dataset, forming the foundation for in-depth analysis. The final tweets were then used as the primary text corpus for conducting topic modeling and analyzing how sentiment (i.e., negative tone and anxiety levels of the tweets calculated using LIWC software) related to different topics changed around major announcements.

3.2 Text Pre-Processing

Prior to conducting the topic modeling analysis, the collected tweets underwent a series of preprocessing steps aimed at enhancing the quality of the data analysis for topic modeling and subsequent analysis (Egger & Yu, 2022). These preprocessing steps were carried out to ensure that the data was in an appropriate format for analysis and to facilitate the ability of BERTopic to effectively capture the underlying semantics and meanings within the text.

The initial step involved text cleaning; wherein various actions were taken to refine the text for analysis. Non-ASCII characters, symbols, digits, and emojis were systematically removed from the text (Dutta et al., 2023). This was done to establish a consistent and uniform textual basis, thereby enabling transformer models to interpret the textual content more accurately (Devlin et al., 2018; Reimers et al., 2019; Medvecki et al., 2024). Importantly, the original casing of words was preserved, as this decision was made to maintain the inherent grammatical structures and nuances present in the text (Reimers et al. 2019; Medvecki et al., 2024). Moreover, the conventional approach of converting all tokens to lowercase was intentionally omitted. This decision was made to preserve crucial information associated with the capitalization of words and retain as much as textual information as possible. Moreover, we opted to retain stop words – commonly occurring words like “and”, “the”, etc. – and not perform lemmatization – reducing words to their base or root form – (Sarica & Luo, 2021; Dutta et al., 2023). Studies have shown that retaining stop words and original word forms can help preserve semantic content and contextual information, which is crucial for capturing the specific ways in which these words contribute to the overall meaning of the text (Sarica & Luo, 2021; Mikolov et al., 2013; Devlin et al., 2018).

The text preprocessing steps undertaken before the BERTopic modeling analysis encompassed the removal of non-essential elements while preserving the natural structure of the text. This approach aimed to empower subsequent analyses, particularly those involving transformer models as these models are capable of a comprehensive and accurate understanding of the semantic context present in the tweets (Vaswani et al., 2017; Devlin et al., 2018; Reimers & Gurevych, 2019, 2020). Next, we discuss the relevance of BERTopic in the context of this study.

3.3 BERTopic

BERTopic, a topic modeling approach which leverages advanced language models like BERT to comprehend words in context (Egger & Yu, 2022; Tang et al., 2024; Williams et al., 2023; Ruocco et al., 2024). This characteristic enables BERTopic to capture complex relationships between words, resulting in more accurate topic identification within text data. BERTopic has demonstrated superior performance compared to traditional methods in tasks involving document grouping and topic labeling (Röder et al., 2015; Lau et al., 2016; Egger & Yu, 2022).

An advantageous feature of BERTopic is its hierarchical topic creation, which organizes topics into sub-topics, providing a more detailed and organized representation of diverse themes in the data (Grootendors, 2022). Yet another significant advantage is transfer learning. Given BERT’s comprehensive pre-training on extensive data, BERTopic can leverage this knowledge to understand new topics without requiring an extensive amount of new data (Grootendors, 2022). While traditional topic models (i.e., LDA and NMF) struggle with short and noisy texts, BERTopic excels in effectively handling such cases (Egger & Yu, 2022). Numerous studies have concluded that BERTopic outperforms LDA and NMF in terms of coherence and topic separation (Egger & Yu, 2022; Röder et al., 2015; Lau et al., 2016; Sánchez-Franco & Rey-Moreno, 2021). This implies that the topics generated by BERTopic are more comprehensible and logical compared to the other models. Furthermore, BERTopic has a high degree of flexibility, adapting well to various data types, documents of varying lengths, and even different languages (Devlin et al., 2018; Wang et al., 2018). This adaptability empowers BERTopic as a potent tool for topic modeling, particularly when dealing with diverse and short text datasets.

The BERTopic approach entails three stages as described in the work of Grootendorst (2022): document embedding (i.e., turning documents into numerical representations), document clustering, and document c-TF-IDF. The initial step involves transforming the corpus documents into word embeddings, a process that maps words into numerical vector spaces while considering semantic nuances and context (Mikolov et al., 2013). In the second step, prior to clustering, embeddings undergo dimensionality reduction, which offers benefits such as eliminating redundant features, reducing computation and training time, and aiding data visualization. The final step involves looking at the importance of words in each group (or cluster), which then helps to distinguish the different topics. This is done using a technique called class-based TF-IDF (or c-TF-IDF). This method measures how important a word is to a whole group/cluster of documents, rather than just one. It helps to create a better picture of what each topic is about (Grootendorst, 2022). The aim of the class-based TF-IDF is to provide the same class vector to all documents within a single class. This regularization reduces the influence of frequently occurring words within a class. The frequency of each word is extracted for each class and divided by the total number of words. This approach models the significance of words in clusters rather than individual documents (Grootendorst, 2022).

3.4 BERTopic Model Implementation

The BERTopic model implementation involves a series of steps that transform textual data into a format suitable for further processing (refer to Fig. 2). This process initiates with the conversion of textual documents into sentence embeddings. In our case, we utilize Sentence-BERT (SBERT) to transform the preprocessed tweets into numerical representation of dimension 384. SBERT is a specialized model designed to provide single embeddings for sentences which focuses on semantic similarity, a crucial aspect of our analysis (Reimers & Gurevych, 2019).

To effectively perform clustering, we recognize the necessity of dimensionality reduction. Many clustering algorithms, including HDBSCAN, are sensitive to the curse of dimensionality, where high-dimensional data points become equidistant which in turn results in poor clustering (Allaoui et al., 2020; Campello et al., 2013). Hence, we employed the Uniform Manifold Approximation and Projection (UMAP) technique to reduce the dimensions of the sentence embeddings. UMAP serves as a dimensionality reduction tool similar to PCA and t-SNE, preserving the essential structural properties of the data (McInnes & Healy, 2018; Allaoui et al., 2020).

Following dimensionality reduction, we used the reduced embedding vectors for clustering. We used HDBSCAN, a density-based algorithm (Campello et al., 2013). Unlike distance-based algorithms like k-means, HDBSCAN considers the reachability of points to form clusters (Campello et al., 2013). Points that are distant from most others are identified as outliers, aiding in the creation of meaningful clusters. While implementing HDBSCAN we specified cluster to be at least 1000 tweets. The training phase of BERTopic model encompassed a combination of above steps which serves as pipeline models. This involved utilizing the generated sentence embeddings of tweets, UMAP model, and HDBSCAN model. The training process yielded the generation of topics, size of topics and representation of every topic (i.e., top terms of every topic).

To enhance topic comprehension, we employed BERTopic's hierarchy visualization to plot the topic hierarchy using the trained BERTopic model and guide the merging of similar topics together (Refer to Section A2.2 of Online Appendix 2). Upon finalizing the merged and updated topic labels (i.e., stage 2 topic allocation) (refer to Table A2.3), we adopted this topic information as our reference point for next section. Refer to Fig. 3 for more information on the topic allocation stages. Next section will discuss about our progression in extending the scope to assign multiple topic labels to a single tweet and discuss the effectiveness of BERTopic model.

3.5 Analyzing BERTopic Model Effectiveness

BERTopic model has a remarkable ability to capture and uncover complex relationships present in textual content (Grootendorst, 2022). We explored its effectiveness to identify prevalent themes within Twitter conversations surrounding small and family-owned businesses.

BERTopic operates by clustering similar tweets into cohesive themes, providing a coherent structure to a very diverse and vast collection of textual content of social media. It uses BERT-based contextual embeddings, which allows BERTopic to understand the text better and have an in-depth understanding of the nuanced semantic connections among words and phrases, which gives it an edge in identifying thematic patterns (Devlin et al., 2018; Reimers & Gurevych, 2019). However, the initial application of BERTopic revealed that a substantial portion of tweets remained classified as outliers, thereby questioning the model's capacity to effectively understand the entire conversational spectrum. In stage 1, BERTopic uncovered a total of 18 meaningful topics and one outlier topic which has tweets which does not fit to the 18 meaningful topics (refer to Table A2.1 in Online Appendix 2). These 18 meaningful topics cover a substantial portion of the tweets — 60,836 out of a grand total of 132,567. However, there remained a vast collection of 71,731 tweets that BERTopic classified as outliers — tweets that didn't fit the themes of the 18 topics identified by BERTopic. Similarly, the number of outliers will not change in stage 2 since we are only merging similar topics other than outlier topic (refer to Table A2.3).

To reduce these outliers, we employed an approach involving cosine similarity. By calculating the average

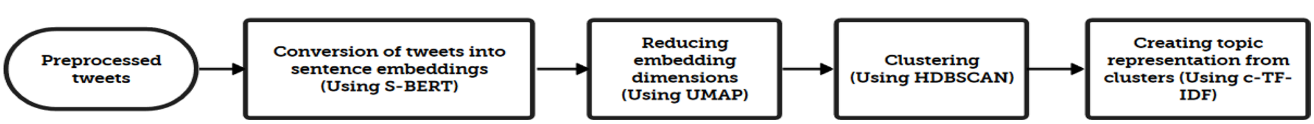

Fig. 2 BERTopic implementation stages

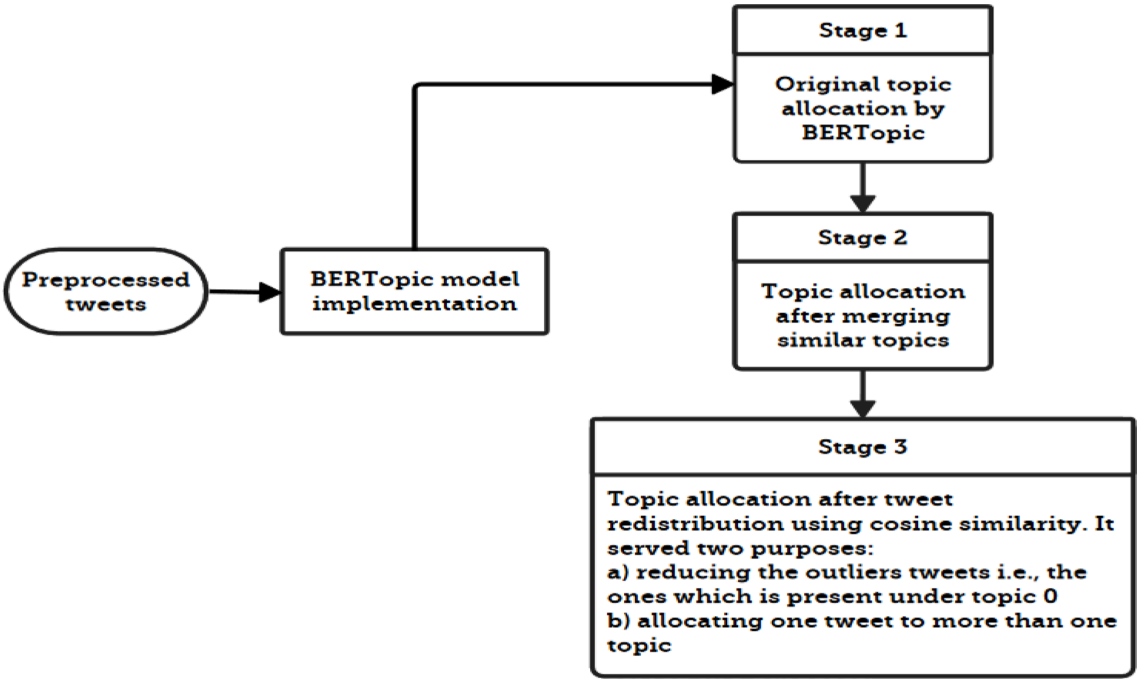

Fig. 3 Topic allocation stages

cosine similarity of a tweet with others tweets in a specific topic, we assessed its affinity with that topic. If this average similarity surpassed the 50% threshold, we added the corresponding topic label to the tweet. In addition to handling outliers, this method also lets tweets be part of more than one topic, depending on how much they were similar with a specific topic. For example, consider tweet_100. Assuming, according to the BERTopic this tweet was initially assigned to topic_7. However, upon calculating the average similarity between tweet_100 and the tweets within topic_1 exceeded 50%, we took the step of modifying the label of tweet_100. Specifically, we changed its label to consider both topic_7 and topic_1. By implementing this approach, our intention was twofold: firstly, to ensure that outlier tweets found their suitable thematic topic, and secondly, to enable a tweet to be associated with multiple topics, guided by the underlying cosine similarities between them. The idea behind the second intention is that there is a possibility that one tweet might be discussing multiple topics. This additional layer of topic assignment was aimed to provide a more nuanced and comprehensive understanding of the themes within the tweets. Please refer to Table A2.4 to see the updated tweet count for each topic and top terms of the topics after implementing this approach. This technique is employed to determine if these outliers could be more appropriately accommodated within the themes identified by the BERTopic. Yet, despite this approach, the increase in tweet numbers per topic was very low, in comparison to Stage 2 (refer to Table A2.5). Note that the stage 2 did not involve outlier reduction as it only merged the similar topics except outlier topic.

After looking into the variation in tweet counts between Stage 3 (with outlier reduction) and Stage 2 (without outlier reduction), it is evident that the cosine similarity approach did not gave a substantial improvement in the tweet count per topic (refer to Table A2.5). However, it is important to note that a substantial increase of 4844 in the tweet count was observed for Topic 11, signifying the potential effectiveness of the introduced approach for certain topics. For subsequent analysis in this study, we have considered the topic allocation data of stage 3.

This discussion highlights that although adding cosine similarity to BERTopic didn't consistently result in a notable increase in the number of tweets per topic, it did show potential in one case, which is Topic 11. Furthermore, this underscores the effectiveness of BERTopic in identifying meaningful and coherent topics. Having said that, it is important to note that a bigger chunk of text (such as blog posts) can exhibit a multifaceted conversation which might not exclusively belong to one topic (Medvecki et al., 2024). Which is why the cosine similarity technique can prove beneficial depending on the nature and length of the text. On platforms like Twitter, where text length tends to be shorter (with an average of 32 words per tweet in our dataset), which reduces the scope of a tweet to be discussing two separate topics or themes. This is a contributing factor to BERTopic's robust performance with Twitter text. Even with our modified approach of assigning tweets to multiple topics and utilizing cosine similarity to reduce outliers, there isn't a substantial increase in the number of tweets per topic. This underscores BERTopic's adaptability and effectiveness in the context of Twitter conversations. Further exploration and fine-tuning of such approaches hold potential for unveiling more comprehensive insights into the dynamic landscape of Twitter discussions on small and family-owned businesses.

4 Results

4.1 Prevalent Topics in Twitter Discussions around Small and Family-Owned Business

After merging similar topics (stage 2) and handling outliers (stage 3), we looked at these topics closely to understand the main themes social media users were discussing about smaller and family-owned businesses on Twitter. An overview of the topics is presented in Table 2. As shown in Table 2, a wide range of topics are covered through our BERT topic modeling approach. The emerging topics are: (1) Entrepreneurship Opportunities during pandemic; (2) Restrictions, challenges and lack of moral support; (3) Loan, grants, and relief programs; (4) Vaccine and mask mandates; (5) Women Entrepreneurs; (6) Canadian small business support issues; (7) Seeking Indian government support; (8) Impact of Pandemic on Black-Owned Small Businesses; (9) Impact of Brexit and Covid on UK small businesses; (10) Impact of Whitmer's Policies on Small Businesses; 11) General support for small businesses. Corresponding word clouds figures (Figs. 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 and 14) are also referenced in Table 2.

4.2 Change in Sentiments Around Major Government Announcements

The methodology used to assess the effectiveness and impact of major government announcements on the topics identified by BERTopic (Stage 3 topic allocation) involves a

Table 2 Topic description

Topic number	Topic	Description	Top keywords figure (Stage 3 Word Clouds)
1	Entrepreneurship Opportunities during pandemic	Discussions on the role of new entrepreneurs in economic recovery and the opportunities that arise by the pandemic.	Figure 4
2	Restrictions, challenges and lack of moral support	Discussions on the impact of lockdowns and restrictions on small businesses, the preferential treatment of large corporations, and the struggles and lack of support faced by small businesses during the pandemic.	Figure 5
3	Loan, grants, and relief programs	Discussions on the financial aid, relief programs, and loan schemes available to small businesses during the COVID-19 pandemic.	Figure 6
4	Vaccine and mask mandates	Discussions of views on COVID-19 mandates such as vaccinations and masks	Figure 7
5	Women Entrepreneurs	Discussions on the hurdles faced by the women entrepreneurs.	Figure 8
6	Canadian small business support issues	Issues and concerns over the support provided to the Canadian small businesses	Figure 9
7	Seeking Indian government support	Tweets discussing the impacts of the lockdown on the Indian small businesses and the need for the support from the government.	Figure 10
8	Impact of Pandemic on Black-Owned Small Businesses	Tweets underscoring the lack of support and challenges faced by the black owned businesses.	Figure 11
9	Impact of Brexit and Covid on UK small businesses	Discussions around the effects of the pandemic and Brexit on UK small businesses.	Figure 12
10	Impact of Whitmer's Policies on Small Businesses	Discussions around the criticism directed towards the inadequacy of Whitmer's policies on small businesses during the pandemic.	Figure 13
11	General support for small businesses	This topic involves discussions of general support for small business rather than being more focused discussions like topic 12 where discussions revolve around PPP loans.	Figure 14

Fig. 4 (Topic 1) Entrepreneurship Opportunities during pandemic

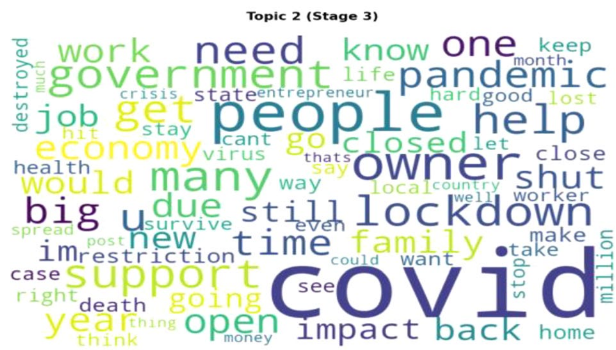

Fig. 5 (Topic 2) Restrictions, challenges and lack of moral support

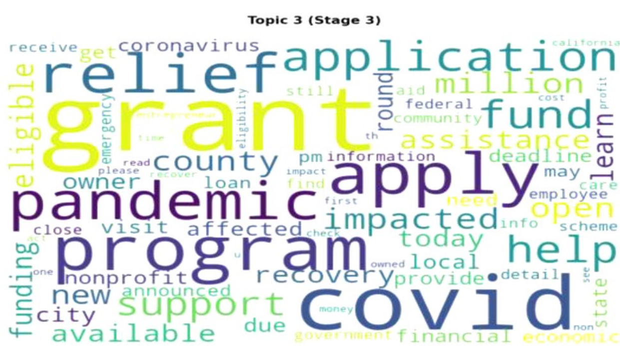

Fig. 6 (Topic 3) Loan, grants, and relief programs

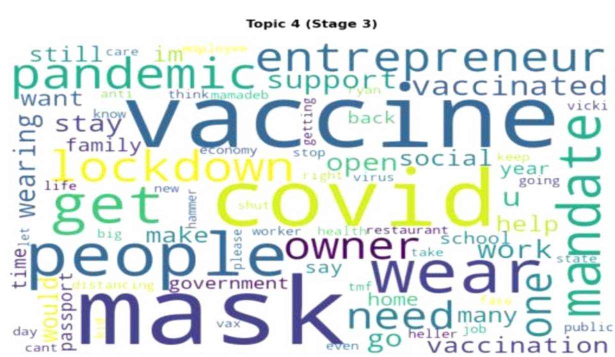

Fig. 7 (Topic 4) Vaccine and mask mandates

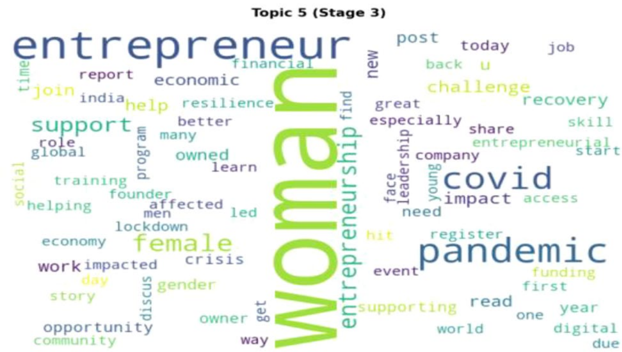

Fig. 8 (Topic 5) Women Entrepreneurs

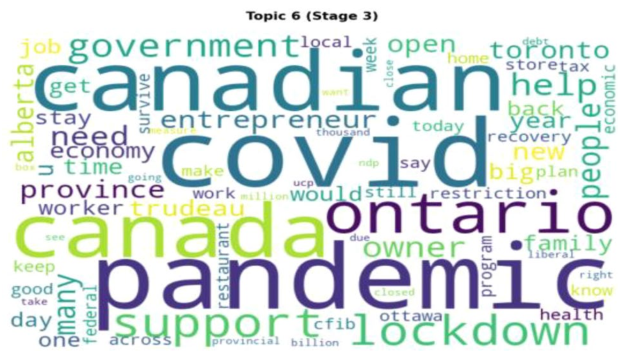

Fig. 9 (Topic 6) Canadian small business support issues

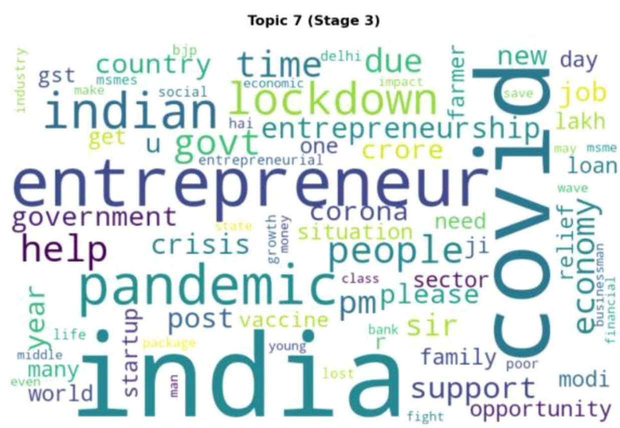

Fig. 10 (Topic 7) Seeking Indian government support

multi-step process. This process systematically collects data on negative tone (tone_neg) and emotional anxiety (emo_anx) using the LIWC software developed by Pennebaker et al. (2007), both before and after the announcements. Here is an in-depth overview of the methodology:

4.2.1 Tweet Collection and Organization

Initially, a collection of tweets related to each BERTopic-identified topic (Stage 3 topic allocation) is gathered. These

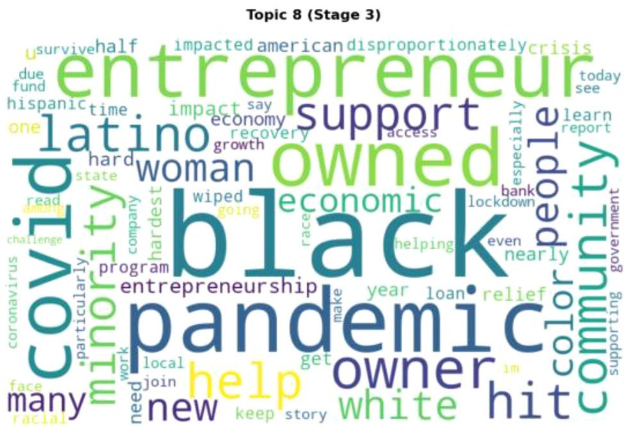

Fig. 11 (Topic 8) Impact of Pandemic on Black-Owned Small Businesses

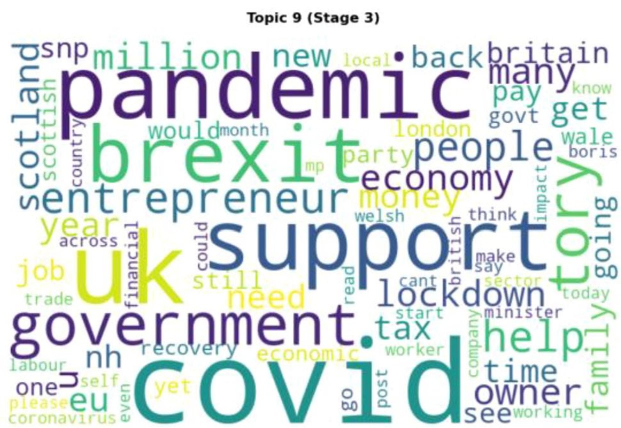

Fig. 12 (Topic 9) Impact of Brexit and Covid on UK small businesses

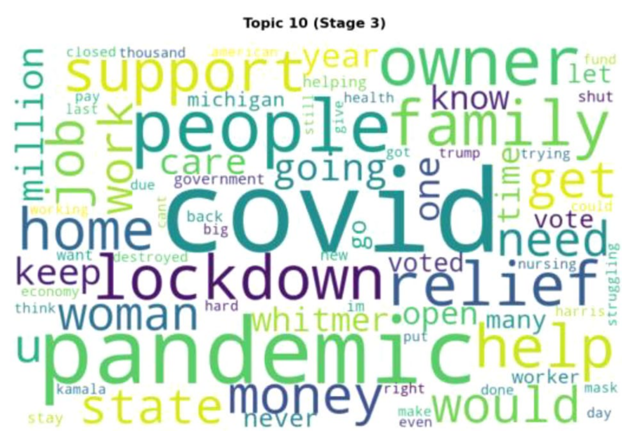

Fig. 13 (Topic 10) Impact of Whitmer's Policies on Small Businesses

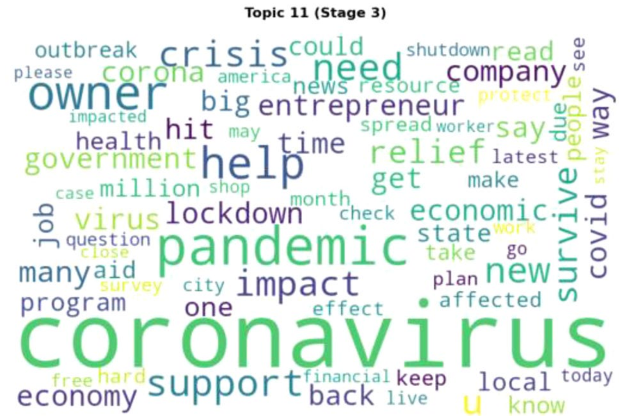

Fig. 14 (Topic 11) General support for small businesses

tweets are systematically organized and stored for subsequent analysis.

4.2.2 Identification of Key Dates and Announcements

Major government announcements that are likely to have an impact on the discussions are identified. Each announcement is associated with a specific date and a brief description outlining the essence of the announcement (refer to Table 3). These announcements serve as focal points for analysis.

4.2.3 Temporal Window Selection

A temporal window of 10 days, 15 days, and 20 days is chosen around each announcement date to understand the effectiveness and impact of the announcements.

4.2.4 Topic-Centric Filtering

For each BERTopic-identified topic (Stage 3 topic allocation), tweets within the chosen temporal window are filtered. These tweets are further categorized based on their topic label, facilitating a focused analysis for each topic.

4.2.5 Collecting LIWC Features

Using LIWC software, we collected the two LIWC features for each tweet related to each topic around the major announcement dates. Specifically, using LIWC software, we collected the features "tone_neg" which measures the extent of negative tone in the tweet, and "emo_anx" which gauges the level of emotional anxiety in the tweet's text. This approach provides an objective view of how negative tone and anxiety levels evolve before and after each major government announcement for the relevant topics. This

Table 3 Key announcements

Announcement	Date	Description
Announcement 1: Fed launched the Paycheck Protection Program Liquidity Facility (PPPLF)	9th April 2020	To help small businesses, the Fed launched the Paycheck Protection Program Liquidity Facility (PPPLF) on April 9, 2020, in concert with the CARES Act. This program lent money to banks so they could, in turn, lend money to small businesses through the Paycheck Protection Program (PPP) (Board of Governors of the Federal Reserve System, 2020, April 9)
Announcement 2: Emergency Rent Subsidy program announced	9th October 2020	The Canada Emergency Rent Subsidy Program (CERS). This program consists of two components: the base rent subsidy, accessible to organizations facing ongoing revenue decline, and the Lockdown Support, which offers supplementary assistance to entities forced to shut down or considerably curtail operations due to public health orders (Department of Finance Canada, 2020, September 10)
Announcement 3: Chancellor Sunak announced a £4.6bn fresh financial support package for struggling	5th January 2021	Chancellor Sunak unveiled a new financial assistance initiative amounting to £4.6 billion, aimed at providing vital support to struggling companies in the UK. This comprehensive package comprises two key components which aims to provide significant financial relief to a wide range of businesses across various sectors, addressing the challenges posed by the ongoing economic disruptions (The Guardian, 2021, January 5)
Announcement 4: SBA increased the maximum loan amount from $150,000 to $500,000	24th March 2021	SBA increased the maximum loan amount from $150,000 to $500,000 (U.S. Small Business Administration, 2021, March 24)
Announcement 5: SBA increased the COVID EIDL maximum loan amount from $500,000 to $2 million	9th September 2021	SBA increased the COVID EIDL maximum loan amount from $500,000 to $2 million (U.S. Small Business Administration, 2021, September 9)

Table 4 Announcement 1 t-test results

		Panel A: 15 days window			Panel B: 20 days window		
topic	liwc_sig	*n* (post)	*n* (pre)	mean_diff (post - pre)	*n* (post)	*n* (pre)	mean_diff (post - pre)
1	tone_neg	103	80	-1.803***	137	89	-1.313**
1	emo_anx	103	80	-0.327*	137	89	-0.257*
2	tone_neg	151	119	0.306	210	158	0.403
2	emo_anx	151	119	-0.058	210	158	-0.079
3	tone_neg	304	202	0.936***	404	249	0.778***
3	emo_anx	304	202	0.074	404	249	0.098**
5	tone_neg	10	8	6.776	11	10	6.419
5	emo_anx	10	8	0	11	10	0
8	tone_neg	9	6	-1.687	19	6	-2.198
8	emo_anx	9	6	0	19	6	0
11	tone_neg	29	43	0.027	45	57	0.697
11	emo_anx	29	43	-0.241	45	57	-0.182

allowed us to gain understanding of the effectiveness of the announcements in alleviating people's concerns.

This comprehensive methodology ensures that the change of sentiment around major government announcements is systematically captured, enabling an objective assessment of the evolving narrative landscape for each BERTopic-identified topic (Stage 3 topic allocation). We use t-test to examine the changes in the negative tone and anxiety levels between pre- and post-announcement periods. To that effect, we primarily used two different time windows (15 days and 20 days) around an announcement[4]. Please note that t-statistics and p-values are not reported in certain cases because the LIWC software did not return any scores for these features. The subsequent section discusses the shifting discourse of Twitter conversations surrounding major announcements, informed by the t-test results.

4.2.6 Evaluation of Announcement Impact

4.2.6.1 Announcement 1: Fed Launched the Paycheck Protection Program Liquidity Facility (A1) From Table 4 Panel A (15-days window), it can be observed that there is significant drop in negative tone and anxiety levels for Topic 1 and significant increase in negative tone for Topic 3 post the announcement. This suggests that the announcement had a positive impact on reducing negativity and anxiety related to Topic 1. In contrast, Topic 3 showed a significant increase in negative tone, but no significant change in emotional

[4] Due to the limited number of observations and insignificant results in the 10-day window, we focused on the 15-day and 20-day windows to examine the changes in negative tone and anxiety levels surrounding the announcement date. The comparison of total observations for each announcement across the 10-day, 15-day, and 20-day windows is presented in Table A2.6, Table A2.7, Table A2.8, Table A2.9, and Table A2.10 in Online Appendix 2.

Table 5 Announcement 2 t-test result

		Panel A: 15 days window			Panel B: 20 days window		
topic	liwc_sig	*n* (post)	*n* (pre)	mean_diff (post - pre)	*n* (post)	*n* (pre)	mean_diff (post - pre)
6	tone_neg	56	31	-0.941	73	40	-1.121*
6	emo_anx	56	31	0.056	73	40	-0.025

Table 6 Announcement 3 t-test results

		Panel A: 15 days window			Panel B: 20 days window		
topic	liwc_sig	*n* (post)	*n* (pre)	mean_diff (post - pre)	*n* (post)	*n* (pre)	mean_diff (post - pre)
9	tone_neg	40	22	0.043	48	28	-0.324
9	emo_anx	40	22	0.128	48	28	0.013

anxiety. This indicates that while the intervention may have influenced the tone of discussions in Topic 3, it did not affect the anxiety levels associated with it. Other topics showed no significant changes in either negative tone or emotional anxiety, implying that the announcement's effects were not widespread across all topics. If we compare these results with Table 4 Panel B (20-days window), the results remain similar except for Topic 3, which sees a significant increase in anxiety levels and the increase in negative tone gets even stronger post-announcement. This suggests that although the announcement on a 15-days window did not affect anxiety levels in Topic 3, its slightly longer term impact (20-days window) was significantly different. In general, these results indicate that Announcement 1 largely failed to alleviate the concerns of people from small and family-owned businesses.

4.2.6.2 Announcement 2: Emergency Rent Subsidy Program Announced (A2) As shown in both panels of Table 5, the results for Topic 6 reveal an interesting trend. In Table 5 Panel A (15-days window), there is no significant change in either negative tone or anxiety. However, in Table 5 Panel B (20-days window), while anxiety remains unchanged, there is a significant decrease in negative tone. This suggests that over the 20-days period, the negative tone has reduced, but anxiety levels remain unaffected. The announcement (A2) shows a degree of effectiveness in reducing negative tone, but only over a longer period, as seen in the 20-days window where a significant decrease in negative tone is observed. Despite this, anxiety levels remain unaffected in both windows, suggesting that while A2 can mitigate negative tone, it did not influenced users' emotional anxiety.

4.2.6.3 Announcement 3: Chancellor Sunak Announced a £4.6bn Fresh Financial Support Package for Struggling (A3) For both Table 6 Panel A (15-days window) and Table 6 Panel B (20-days window), there are no significant changes in either negative tone or anxiety for Topic 6. This indicates that A3 had no effect on Topic 6, and consequently, did not appear to have any significant impact on negative tone or anxiety in either the 15-days or 20-days window. This suggests that the announcement was ineffective in influencing user negative tone or anxiety levels, implying that the announcement of the fresh financial support for the UK small businesses did not resonate or address the concerns related to Topic 6 effectively, failing to mitigate negative tone or anxiety levels.

4.2.6.4 Announcement 4: SBA increased the maximum loan amount from \$150,000 to \$500,000 (A4) For most topics (1, 2, 5, 8, and 11), announcement 4 (A4) does not lead to any significant changes in negative tone or anxiety, indicating a limited scope of effectiveness. However, for Topic 3, the results are more complex. While A4 effectively reduces anxiety in both the 15-days and 20-days windows, positively influencing users' emotional state, its impact on negative tone is inconsistent. There is a significant increase in negative tone in the 15-day window, which does not persist into the 20-day window. This suggests that the announcement might have initially triggered negative reactions that dissipated over time. Comparing Table 7 Panel A (15-days window) results with Table 7 Panel B (20-days window) reveals consistency in results for most topics. However, for Topic 3, the significant increase in negative tone observed over the 15-day period does not persist into the 20-day period, while the significant decrease in anxiety persists across both windows. Overall, A4 presents mixed results, with a positive impact on anxiety for Topic 3, but an inconsistent effect on negative tone, and limited effectiveness for other topics.

4.2.6.5 Announcement 5: SBA increased the COVID EIDL maximum loan amount from \$500,000 to \$2 million (A5) The results from both Table 8 Panel A (15-days window) and Table 8 Panel B (20-days window) show a consistent lack of significant changes in negative tone or anxiety across all topics, indicating that announcement 5 (A5) had

Table 7 Announcement 4 t-test results

		Panel A: 15 days window			Panel B: 20 days window		
topic	liwc_sig	*n* (post)	*n* (pre)	mean_diff (post - pre)	*n* (post)	*n* (pre)	mean_diff (post - pre)
1	tone_neg	192	248	0.136	276	311	0.376
1	emo_anx	192	248	-0.054	276	311	-0.010
2	tone_neg	471	434	-0.068	627	564	-0.026
2	emo_anx	471	434	-0.060	627	564	-0.068
3	tone_neg	211	338	0.451*	252	525	0.252
3	emo_anx	211	338	-0.148***	252	525	-0.105**
5	tone_neg	70	130	0.703	86	223	0.536
5	emo_anx	70	130	0.021	86	223	0.021
8	tone_neg	35	38	1.183	39	50	0.105
8	emo_anx	35	38	0	39	50	0
11	tone_neg	10	15	-0.919	13	20	-1.909
11	emo_anx	10	15	-0.215	13	20	-0.161

Table 8 Announcement 5 t-test results

		Panel A: 15 days window			Panel B: 20 days window		
topic	liwc_sig	*n* (post)	*n* (pre)	mean_diff (post - pre)	*n* (post)	*n* (pre)	mean_diff (post - pre)
1	tone_neg	142	129	-0.442	176	190	-0.371
1	emo_anx	142	129	0.016	176	190	0.071
2	tone_neg	286	252	0.230	362	332	0.293
2	emo_anx	286	252	-0.070	362	332	-0.056
3	tone_neg	94	83	0.318	119	112	0.395
3	emo_anx	94	83	-0.006	119	112	0.032
5	tone_neg	46	28	-0.253	51	37	-0.526
5	emo_anx	46	28	0.063	51	37	0.057
8	tone_neg	16	16	1.945	23	19	1.204
8	emo_anx	16	16	0	23	19	0.248
11	tone_neg	4	2	3.120	6	2	1.630
11	emo_anx	4	2	0.675	6	2	0.450

no effect on these topics. In other words, A5 does not result in any significant changes in negative tone or anxiety across all topics in both the 15-days and 20-days windows. In other words, A5, which pertains to the US Small Business Administration's (SBA) increase in the COVID Economic Injury Disaster Loan (EIDL) maximum loan amount from $500,000 to $2 million, does not result in any significant changes in negative tone or anxiety across all topics in both the 15-day and 20-day windows. Despite the significant increase in loan amount, the announcement appears to have failed to reduce the negative tone and anxiety levels of the users related to the small businesses.

5 Summary of Research Findings

This study centered around two main research questions: (1) What are the small and family-owned businesses topics do social media users discuss during the COVID-19 pandemic? and (2) What is the impact of major government announcements on these discussions?

The first research question focuses on the exploration of various topics discussed by social media users in the context of smaller and family-owned businesses during the COVID-19 period. Interestingly, these discussions are not confined to operational and financial challenges faced by small and family-owned businesses. Instead, a wide range of topics were captured through our topic modeling analysis, such as, entrepreneurship opportunities during the pandemic, conversations surrounding the restrictions, challenges, and lack of moral support, the struggles faced by small businesses, discussions on financial aid, relief programs, and loan schemes, views on vaccine and mask mandates, the hurdles faced by women entrepreneurs, issues concerning the support provided to Canadian small businesses, the lack of support for Black-owned small businesses, and the combined impact of Brexit and COVID-19 on UK small businesses.

The second part of the research explored the effectiveness of government announcements aimed at mitigating the concerns of small businesses during the pandemic, revealing largely ineffective outcomes with a few notable exceptions. Announcement 1, which introduced the Paycheck Protection Program Liquidity Facility (PPPLF) by the US Federal

Reserve, showed some efficacy in reducing negative sentiment and anxiety related to entrepreneurship opportunities (Topic 1). However, its impact was not uniformly felt across all topics and, interestingly, resulted in a significant increase in both negative tone and anxiety levels in discussions related to loans, grants, and relief programs (Topic 3). Announcement 2, unveiling the Emergency Rent Subsidy program in Canada, effectively reduced negative tone concerning Canadian small business support issues (Topic 6) over a 20-days period, though it did not influence anxiety levels. Announcement 4, which saw the US Small Business Administration (SBA) increase the maximum loan amount from $150,000 to $500,000, yielded mixed results. Discussions of loan, grants, and relief programs (Topic 3) showed a reduction in anxiety, but the effects on negative tone were inconsistent.

Conversely, most announcements, including Announcement 3 and Announcement 5, had no significant impact on negative tone or anxiety levels across all topics. Failure of both these announcements to produce any significant impact is particularly noteworthy, given the fresh financial support of £4.6 billion and substantial increase in loan amount; indicating that these announcements did not resonate with or effectively address the needs and concerns of small and family-owned businesses. The lack of impact, especially in critical areas such as general support for small businesses (Topic 11) and specific issues like the impact of pandemic policies on Black-owned small businesses (Topic 8), underscores the need for more targeted and responsive measures to support this vital sector.

6 Policymaking and Practical Implications

From a policymaking perspective, the findings of this study carry essential implications, offering valuable insights into the dynamics of social media discussions during times of crisis. Firstly, it is evident from the findings of the study that policymakers should consider the importance of tailoring their support programs to accommodate the unique challenges and opportunities faced by different demographics, such as female and black entrepreneurs. This approach can ensure that assistance is effective and relevant to the specific circumstances of these small and family-owned businesses. Policymakers should strive to ensure that relief measures are designed with fairness in mind, addressing gender disparities and promoting access to support for marginalized small and family-owned businesses. Moreover, policies developed through consultation with a diverse range of stakeholders can significantly enhance their effectiveness, fostering a more inclusive and supportive environment for all.

The practical implications of this research are closely tied to the real-time nature of social media as a source of information. Policymakers can benefit from social media platforms by gaining immediate insights into the evolving concerns and sentiments of small and family-owned businesses during crises. Further, with the conclusion of the COVID-19 pandemic and the recent alert from the World Health Organization (WHO) regarding the potential emergence of a new pandemic, this study offers a methodological framework for policymakers to harness the wealth of information available on social media. The methodology proposed in this study leverages Natural Language Processing (NLP) techniques, making it a valuable tool for making well-informed and timely policy decisions. The incorporation of BERTopic and relevant statistical analysis offers pragmatic insights for governments and organizations seeking to communicate policy measures effectively. By understanding the alignment between announcements and existing topics, these insights aid in constructing messages that genuinely resonate with the public, ensuring clear communication of policy intentions.

7 Conclusion

During a crisis, small and family-owned businesses tend to experience more severe economic consequences than their larger counterparts and often lack financial resources needed to weather the challenges brought about by the crisis. To comprehend the distinct challenges and concerns of small and family-owned businesses during a major crisis, this research study focuses on the recent COVID-19 pandemic, which had a catastrophic effect on businesses and societies alike. To that effect, we addressed two research questions: First, what topics pertaining to small and family-owned businesses do social media users discuss during the COVID-19 pandemic? To achieve this goal, we employed the BERTopic model, a state-of-the-art technique for topic modeling, to identify and categorize prevalent themes arising from the discourse. Second, what is the impact of major government announcements in terms of changing negative sentiments of the users? Our findings suggest that government announcements often fell short of effectively addressing the concerns and needs of small businesses during the pandemic. This highlights the importance of a better consultation process and communication strategy by policymakers.

Our study has a few limitations that could be addressed in future research. The first limitation pertains to the platform (i.e., Twitter) utilized in our study. Twitter was acquired by Elon Musk on 27th October 2022 (The New York Times, 2022, October 27) which presents some challenges for future Twitter users and researchers. (a) there has been a noticeable reduction in activity among the "Twitter's most active users" and approximately 25% decrease in the monthly average tweets of "top users" (CNBC, 2023, May 19). (b) the blue check mark which previously was assigned to authentic verified users on Twitter now has become a paid feature leading to any user on twitter to acquire the blue check mark which becomes difficult for users and researchers to differentiate information from real and false accounts. This could give rise to dissemination of misinformation (The New York Times, 2023, March 31). Finally, the termination of academic access to the Twitter API, coupled with the introduction of new paid plans in early 2023, has created significant challenges for numerous researchers (CNN, 2023, April 5; CNN, 2023, February 9). Moving forward, the limitations imposed by this shift could impede researchers' ability to access data pertaining to public opinion.

Second, Twitter is just one of many social media platforms. In the future, other social media platforms like Reddit (Ruan & Lv, 2022), Facebook (Smith & Graham, 2017), and Telegram (Vergani et al., 2022) could be utilized to explore similar research questions. Further, while quantitative analysis through topic modeling is valuable, incorporating qualitative insights through interviews or case studies could enrich the understanding of the challenges and strategies of small and family-owned businesses. This could provide depth to the findings and validate or challenge the results from the topic modeling approach. Future studies may explore such methodologies Third, as per the research scope of this study, we focused on the small and family-owned business exclusively. Future studies may perform some comparative analysis and show how discussion topics differ between 'large' and 'small and family-owned firms'.

Finally, our study exclusively focused on the impact of COVID-19 pandemic; however, there are other significant yet short-time crises that may entail different discussion topics, which could be studied in the future. Unlike a major pandemic such as COVID-19, short-term crises may require different types of interventions from governments or policymakers compared to prolonged crises. Despite this limitation, we argue that the findings of our study remain relevant for future crises in two significant ways. (a) the implications of our study extend beyond the immediate effects of COVID-19. The World Health Organization (WHO) warns that even more severe health and socio-economic crises may occur in the future, necessitating better preparedness for handling subsequent devastating effects. (b) as discussed in the introduction, small businesses face considerable challenges in any crisis. These businesses often struggle due to inherent vulnerabilities such as inadequate planning, restricted cash flow, limited access to capital, lack of government assistance, and infrastructure deficiencies (Runyan, 2006). Therefore, while our study's generalizability may be constrained to major, enduring disasters, its insights are still applicable to future health and socio-economic crises, offering valuable lessons for improving crisis preparedness and response strategies.

Acknowledgements N/A.

Author contributions All authors made significant contributions towards the development of this manuscript. For instance, all authors contributed to the study conception and design, material preparation, data collection and analysis. The first draft of the manuscript was developed jointly, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding No funding was received to assist with the preparation of this manuscript.

Data Availability As outlined in the article, the name of the secondary data source (used in this study) is revealed, and any interested researcher can replicate the process.

Declarations

Ethics approval and consent to participate This study is based on secondary data, for which we did not require any ethical approval.

Consent for publication If the study is accepted, we give our consent for publication.

Competing interests All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

References

Ackah, C., Görg, H., Hanley, A., & Hornok, C. (2023). Africa's businesswomen – underfunded or underperforming? *Small Business Economics*, *62*(3), 1051–1074.

Alekseev, G., Amer, S., Gopal, M., Kuchler, T., Schneider, J. W., Stroebel, J., & Wernerfelt, N. (2023). The effects of COVID-19 on U.S. small businesses: Evidence from owners, managers, and employees. *Management Science*, *69*(1), 7–24.

Allaoui, M., Kherfi, M. L., & Cheriet, A. (2020). Considerably improving clustering algorithms using UMAP Dimensionality Reduction Technique: A Comparative study. In *Lecture notes in computer science* (pp. 317–325).

Asghari, M., Sierra-Sosa, D., & Elmaghraby, A. (2018). Trends on health in social media: Analysis using Twitter topic modeling. In *2018 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)* (pp. 558–563). IEEE.

Bachura, E., Valecha, R., Chen, R., & Rao, H. R. (2022). The OPM data breach: An investigation of shared emotional reactions on Twitter. *Management Information Systems Quarterly*, *46*(2), 881–910.

Bartik, A., Bertrand, M., Cullen, Z., Glaeser, E. L., Luca, M., & Stanton, C. (2020). How are small businesses adjusting to COVID-19? Early evidence from a survey. *HKS Working Paper No. RWP20-012, University of Chicago, Becker Friedman Institute for Economics Working Paper No. 2020-42*.

Belghitar, Y., Moro, A., & Radić, N. (2021). When the rainy day is the worst hurricane ever: The effects of governmental policies on SMEs during COVID-19. *Small Business Economics*, *58*(2), 943–961.

Belitski, M., Guenther, C., Kritikos, A. S., & Thurik, R. (2021). Economic effects of the COVID-19 pandemic on entrepreneurship and small businesses. *Small Business Economics*, *58*(2), 593–609.

Bhatt, P., Vemprala, N., Valecha, R., Hariharan, G., & Rao, H. R. (2022). User privacy, Surveillance and Public Health during COVID-19 – an examination of Twitterverse. *Information Systems Frontiers*, *25*(5), 1667–1682.

Board of Governors of the Federal Reserve System (2020, April 9). Federal Reserve takes additional actions to provide up to $2.3 trillion in loans to support the economy. Retrieved from https://www.federalreserve.gov/newsevents/pressreleases/monetary20200409a.htm Accessed on 2024, April 25.

Bruhn, M., Demirguc-Kunt, A., & Singer, D. (2023). Competition and firm recovery post-COVID-19. *Small Business Economics*, *61*, 1555–1586.

Brülhart, M., Lalive, R., Lehmann, T., & Siegenthaler, M. (2020). COVID-19 financial support to small businesses in Switzerland: Evaluation and outlook. *Swiss Journal of Economics and Statistics*, *156*(1).

Campello, R. J. G. B., Moulavi, D., & Sander, J. (2013). Density-Based clustering based on hierarchical density estimates. In *Lecture notes in computer science* (pp. 160–172).

Chang, H. H., & Chuang, S. (2011). Social capital and individual motivations on knowledge sharing: Participant involvement as a moderator. *Information & Management*, *48*(1), 9–18.

Chen, E., Lerman, K., & Ferrara, E. (2020). Tracking social media discourse about the COVID-19 pandemic: Development of a public coronavirus Twitter data set. *JMIR Public Health and Surveillance*, 6(2), e19273.

Chhatwani, M., Mishra, S. K., Varma, A., & Rai, H. (2022). Psychological resilience and business survival chances: A study of small firms in the USA during COVID-19. *Journal of Business Research*, *142*, 277–286.

Cho, H., Li, P., Ngien, A., Tan, M. G., Chen, A., & Nekmat, E. (2023). The bright and dark sides of social media use during COVID-19 lockdown: Contrasting social media effects through social liability vs. social support. *Computers in Human Behavior*, *146*, 107795.

Choi, D., & Kim, P. (2013). Sentiment analysis for tracking breaking events: a case study on Twitter. In Lecture notes in computer science (pp. 285–294).

Choudrie, J., Patil, S., Kotecha, K., Matta, N., & Pappas, I. O. (2021). Applying and understanding an advanced, novel deep learning approach: A COVID 19, text based, emotions analysis study. *Information Systems Frontiers*, *23*(6), 1431–1465.

Chu, S., Chen, H., & Sung, Y. (2015). Following brands on Twitter: An extension of theory of planned behavior. *International Journal of Advertising*, *35*(3), 421–437.

CNBC. (2023, May 19). Twitter's top users are posting less since Musk takeover last year, Pew Survey shows. Retrieved from https://www.cnbc.com/2023/05/17/twitters-top-users-are-posting-less-since-musk-acquisition-pew-shows.html Accessed on 2023, November 14.

CNN. (2023, February 9). Elon Musk pledged transparency at Twitter. But he's walling off researchers. Retrieved from https://www.cnn.com/2023/02/09/tech/musk-twitter-transparency-researchers/index.html Accessed on 2023, November 14.

CNN. (2023, April 5). Academic researchers blast Twitter's data paywall as 'outrageously expensive.'. Retrieved from https://www.cnn.com/2023/04/05/tech/academic-researchers-blast-twitter-paywall/index.html Accessed on 2023, November 14.

Corti, L., Zanetti, M., Tricella, G., & Bonati, M. (2022). Social media analysis of Twitter tweets related to ASD in 2019–2020, with particular attention to COVID-19: Topic modelling and sentiment analysis. *Journal of Big Data*, *9*(1).

Cowling, M., Liu, W., & Calabrese, R. (2021). Has previous loan rejection scarred firms from applying for loans during Covid-19? *Small Business Economics*, *59*(4), 1327–1350.

Crannell, W. C., Clark, E. M., Jones, C., James, T. A., & Moore, J. (2016). A pattern-matched Twitter analysis of US cancer-patient sentiments. *Journal of Surgical Research*, *206*(2), 536–542.

D'Andrea, E., Ducange, P., Bechini, A., Renda, A., & Marcelloni, F. (2019). Monitoring the public opinion about the vaccination topic from tweets analysis. *Expert Systems with Applications*, *116*, 209–226.

Department of Finance Canada (2020, September 10). Government introduces the new Canada Emergency Rent Subsidy program. Retrieved from https://www.canada.ca/en/department-finance/news/2020/11/canada-emergency-rent-subsidy.html Accessed on 2024, April 25.

Devlin, J., Chang, M., Lee, K., & Toutanova, K. (2018). *BERT: Pre-training of Deep Bidirectional transformers for Language understanding*. arXiv (Cornell University.

Doern, R. (2014). Entrepreneurship and crisis management: The experiences of small businesses during the London 2011 riots. *International Small Business Journal*, *34*(3), 276–302.

Dörr, J. O., Licht, G., & Murmann, S. (2021). Small firms and the COVID-19 insolvency gap. *Small Business Economics*, *58*(2), 887–917.

Dutta, S., Kumar, A., Pant, P., Walsh, C., & Dutta, M. (2023). Using 10-K text to gauge COVID-related corporate disclosure. *PloS One*, 18(3), e0283138.

Dy, A. M., & Jayawarna, D. (2020). *Bios, mythoi* and women entreprenAuwynterianterian analysis of the intersectional impacts of the COVID-19 pandemic on self-employed women and women-owned businesses. *International Small Business Journal*, *38*(5), 391–403.

Egger, R., & Yu, J. (2022). A topic modeling comparison between LDA, NMF, Top2VEC, and BERTopic to demystify Twitter posts. *Frontiers in Sociology*, *7*.

Fairlie, R. W., & Fossen, F. M. (2021). The early impacts of the COVID-19 pandemic on business sales. *Small Business Economics*, *58*(4), 1853–1864.

FamilyBusiness (2021, June 2). Measuring the Financial Impact of Family Businesses on the US Economy. Retrieved from https://familybusiness.org/content/measuring-the-financial-impact-of-family-businesses-on-the-US-ec Accessed on 2023, November 14.

Giachanou, A., & Crestani, F. (2016). Like it or not. *ACM Computing Surveys*, *49*(2), 1–41.

Golbeck, J., Grimes, J. M., & Rogers, A. (2010). Twitter use by the U.S. Congress. *Journal of the American Society for Information Science and Technology*, *61*(8), 1612–1621.

González-Padilla, D. A., & Tortolero-Blanco, L. (2020). Social media influence in the COVID-19 pandemic. *International Braz J Urol*, *46*(suppl 1), 120–124.

Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. arXiv (Cornell University).

Hadjielias, E., Christofi, M., & Tarba, S. (2022). Contextualizing small business resilience during the COVID-19 pandemic: Evidence from small business owner-managers. *Small Business Economics*, *59*(4), 1351–1380.

Hidalgo, C. R., Tan, E., & Verlegh, P. (2015). The social sharing of emotion (SSE) in online social networks: A case study in Live Journal. *Computers in Human Behavior*, *52*, 364–372.

Irvine, W., & Anderson, A. R. (2006). The impacts of foot and mouth disease on a peripheral tourism area: The role and effect of crisis

management. *Journal of Travel & Tourism Marketing*, *19*(2–3), 47–60. https://doi.org/10.1300/j073v19n02_05
Kalogiannidis, S. (2020). Covid impact on small business. *International Journal of Social Science and Economics Invention*, *6*(12).
Khatua, A., Khatua, A., & Wang, Z. (2020). Predicting political sentiments of voters from Twitter in multi-party contexts. *Applied Soft Computing*, *97*, 106743.
Kritikos, A. S., Graeber, D., & Seebauer, J. (2020). Pandemie wird zur Krise für Selbständige. DIW aktuell 47.
Lal, B., Dwivedi, Y. K., & Haag, M. (2021). Working from Home during Covid-19: Doing and managing Technology-enabled Social Interaction with colleagues at a Distance. *Information Systems Frontiers*, *25*(4), 1333–1350.
Lau, J., & Baldwin, T. (2016). An Empirical Evaluation of doc2vec with Practical Insights into Document Embedding Generation. In *Proceedings of the 1st Workshop on Representation Learning for NLP* (pp. 78–86). Association for Computational Linguistics.
Lee, J., Rehman, B. A., Agrawal, M., & Rao, H. R. (2016). Sentiment analysis of Twitter users over time: The case of the Boston bombing tragedy. In Lecture notes in business information processing (pp. 1–14).
Lin, T., Zhang, X., Wang, Y., & Liu, H. (2020). Early detection of rumours on Twitter via stance transfer learning. In *Lecture Notes in Computer Science* (pp. 575–588).
Lips, A. (2021). The situation of young people at home during COVID-19 pandemic. *Childhood Vulnerability Journal/Childhood Vulnerability Journal*, *3*(1–3), 61–78.
Luminet, O., Bouts, P., Delie, F., Manstead, A. S. R., & Rimé, B. (2000). Social sharing of emotion following exposure to a negatively valenced situation. *Cognition and Emotion*, *14*(5), 661–688.
Manolova, T. S., Brush, C. G., Edelman, L. F., & Elam, A. B. (2020). Pivoting to stay the course: How women entrepreneurs take advantage of opportunities created by the COVID-19 pandemic. *International Small Business Journal*, *38*(6), 481–491.
Marshall, M. I., Niehm, L. S., Sydnor, S., & Schrank, H. L. (2015). Predicting small business demise after a natural disaster: An analysis of pre-existing conditions. *Natural Hazards*, *79*(1), 331–354.
McInnes, L., & Healy, J. (2018). UMAP: Uniform manifold approximation and projection for dimension reduction. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1802.03426
McKibbin, W., & Fernando, R. (2020). *Economics in the time of COVID-19, (*45–53), CEPR.
McNamara, K. E. M. (2013). A state of emergency: How local businesses experienced the 2012 flood in Fiji. *The Australian Journal of Emergency Management*, *28*(3), 17–23.
Medvecki, D., Bašaragin, B., Ljajić, A., & Milošević, N. (2024). Multilingual Transformer and BERTopic for short text topic modeling: The case of Serbian. In *Lecture notes in networks and systems* (pp. 161–173).
Mejova, Y., Weber, I., & Fernandez-Luque, L. (2018). Online Health monitoring using Facebook Advertisement Audience Estimates in the United States: Evaluation study. *JMIR Public Health and Surveillance*, 4(1), e30.
Mendon, S., Dutta, P., Behl, A., & Lessmann, S. (2021). A Hybrid Approach of machine learning and lexicons to sentiment analysis: Enhanced insights from Twitter Data of Natural disasters. *Information Systems Frontiers*, *23*(5), 1145–1168.
Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed Representations of Words and Phrases and their Compositionality. *arXiv (Cornell University)*.
Onyeaka, H., Anumudu, C., Al-Sharify, Z. T., Egele-Godswill, E., & Mbaegbu, P. (2021). COVID-19 pandemic: A review of the global lockdown and its far-reaching effects. *Science Progress*, *104*(2), 003685042110198.
Panahi, S., Watson, J., & Partridge, H. (2015). Information encountering on social media and tacit knowledge sharing. *Journal of Information Science*, *42*(4), 539–550.
Panahi, S., Watson, J., & Partridge, H. (2016). Conceptualising social media support for tacit knowledge sharing: Physicians' perspectives and experiences. *Journal of Knowledge Management*, *20*(2), 344–363.
Papadopoulos, T., Stamati, T., & Nopparuch, P. (2013). Exploring the determinants of knowledge sharing via employee weblogs. *International Journal of Information Management*, *33*(1), 133–146.
Patel, P. C., & Rietveld, C. A. (2020). The impact of financial insecurity on the self-employed's short-term psychological distress: Evidence from the COVID-19 pandemic. *Journal of Business Venturing Insights*, *14*, e00206.
Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). *Linguistic inquiry and word count: LIWC [Computer software]*. LIWC.net.
Pollack, C., Gilbert-Diamond, D., Onega, T., Vosoughi, S., O'Malley, A. J., & Emond, J. A. (2023). Obesity-related discourse on Facebook and Instagram throughout the COVID-19 pandemic: Comparative longitudinal evaluation. *JMIR Infodemiology*, *3*, e40005.
Rahmadan, M. C., Hidayanto, A. N., Ekasari, D. S., Purwandari, B., & Theresiawati (2020). Sentiment analysis and topic modelling using the LDA method related to the flood disaster in Jakarta on Twitter. In *2020 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)* (pp. 126–130). IEEE.
Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using Siamese BERT-networks. In K. Inui, J. Jiang, V. Ng, & X. Wan (Eds.), *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* (pp. 3982–3992). Association for Computational Linguistics.
Reimers, N., & Gurevych, I. (2020). *Making Monolingual Sentence Embeddings Multilingual using knowledge distillation*. arXiv (Cornell University.
Reuters (2023, August 18). US CDC tracks new lineage of virus that causes COVID. Retrieved from https://www.reuters.com/world/us/us-cdc-tracks-new-lineage-virus-that-causes-covid-2023-08-18 Accessed on 2023, November 14.
Rimé, B. (2009). Emotion elicits the social sharing of emotion: Theory and empirical review. *Emotion Review*, *1*(1), 60–85.
Rimé, B., Finkenauer, C., Luminet, O., Zech, E., & Philippot, P. (1998). Social sharing of emotion: New evidence and new questions. *European Review of Social Psychology*, *9*(1), 145–189.
Röder, M., Both, A., & Hinneburg, A. (2015). Exploring the Space of Topic Coherence Measures. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining* (pp. 399–408). Association for Computing Machinery.
Ruan, T., & Lv, Q. (2022). Public perception of electric vehicles on reddit over the past decade. *Communications in Transportation Research*, *2*, 100070.
Runyan, R. C. (2006). Small Business in the Face of Crisis: Identifying barriers to recovery from a natural Disaster1. *Journal of Contingencies and Crisis Management*, *14*(1), 12–26.
Ruocco, L., Zhuang, Y., Ng, R. T., Munthali, R. J., Hudec, K. L., Wang, A. Y., Vereschagin, M., & Vigo, D. (2024). A platform for connecting social media data to domain-specific topics using large language models: an application to student mental health. *JAMIA Open*, 7(1).
Şahin, A., Kitao, S., Cororaton, A., & Laiu, S. (2011). Why small businesses were hit Harder by the recent recession. *Current Issues in Economics and Finance*, *17*(14), 1–7.
Sánchez-Franco, M. J., & Rey-Moreno, M. (2021). Do travelers' reviews depend on the destination? An analysis in coastal and

urban peer-to-peer lodgings. *Psychology & Marketing*, *39*(2), 441–459.
Sarica, S., & Luo, J. (2021). Stopwords in technical language processing. *PloS One*, *16*(8), e0254937.
Scherer, K. R. (2005). What are emotions? And how can they be measured? *Information Sur Les Sciences Sociales/Social Science Information*, *44*(4), 695–729.
Small Business Administration, U. S. (2021, March 24). SBA to increase lending limit for COVID-19 Economic Injury Disaster Loans. Retrieved from https://www.sba.gov/article/2021/mar/24/sba-increase-lending-limit-covid-19-economic-injury-disaster-loans Accessed on 2024, April 25.
Small Business Administration, U. S. (2021, September 9). COVID-19 Economic Injury Disaster Loan (EIDL). Retrieved from https://www.sba.gov/funding-programs/loans/covid-19-relief-options/covid-19-economic-injury-disaster-loan/about-covid-19-eidl Accessed on 2024, April 25.
Smith, N., & Graham, T. (2017). Mapping the anti-vaccination movement on Facebook. *Information Communication & Society*, *22*(9), 1310–1327.
Soluk, J., Kammerlander, N., & De Massis, A. (2021). Exogenous shocks and the adaptive capacity of family firms: Exploring behavioral changes and digital technologies in the COVID-19 pandemic. *R&D Management*, *51*(4), 364–380.
Steinert, S. (2020). Corona and value change. The role of social media and emotional contagion. *Ethics and Information Technology*, *23*(S1), 59–68.
Stephens, S. D. G., Cunningham, I., & Kabir, Y. (2021). Female entrepreneurs in a time of crisis: Evidence from Ireland. *International Journal of Gender and Entrepreneurship*, *13*(2), 106–120.
Tang, J., Zhang, S. X., & Lin, S. (2021). To reopen or not to reopen? How entrepreneurial alertness influences small business reopening after the COVID-19 lockdown. *Journal of Business Venturing Insights*, *16*, e00275.
Tang, Z., Pan, X., & Gu, Z. (2024). Analyzing public demands on China's online government inquiry platform: A BERTopic-Based topic modeling study. *PloS One*, 19(2), e0296855.
The New York Times (2022, October 27). Elon Musk Completes $44 Billion Deal to Own Twitter. Retrieved from https://www.nytimes.com/2022/10/27/technology/elon-musk-twitter-deal-complete.html Accessed on 2023, November 14.
The Guardian (2021, January 5). Rishi Sunak unveils £4.6bn relief package for UK retail and hospitality sectors. The Guardian. Retrieved from https://www.theguardian.com/politics/2021/jan/05/rishi-sunak-unveils-46bn-relief-package-for-uk-retail-and-hospitality-sectors Accessed on 2024, April 25.
The New York Times (2023, March 31). Twitter's Blue Check Apocalypse Is Upon Us. Here's What to Know. Accessed on 2023, November 14.
Tiwari, P., Yadav, P., Kumar, S., Mishra, B. K., Nguyen, G. N., Gochhayat, S. P., Singh, J., & Prasad, M. (2019). Sentiment analysis for airlines services based on Twitter Dataset. In *Elsevier eBooks* (pp. 149–162). https://doi.org/10.1016/b978-0-12-815458-8.00008-6
Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *arXiv (Cornell University)*.
Vergani, M., Arranz, A. M., Scrivens, R., & Orellana, L. (2022). Hate speech in a Telegram conspiracy channel during the first year of the COVID-19 pandemic. *Social Media + Society*, *8*(4), 205630512211387.
Wang, Y., & Yang, Y. (2020). Dialogic communication on social media: How organizations use Twitter to build dialogic relationships with their publics. *Computers in Human Behavior*, *104*, 106183.
Wang, A., Singh, A., Michael, J., Hill, F., Levy, O., & Bowman, S. (2018). GLUE: A Multi-Task Benchmark and Analysis Platform for Natural Language Understanding. In *Proceedings of the 2018 EMNLP Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP* (pp. 353–355). Association for Computational Linguistics.
Wedawatta, G., Ingirige, B., & Proverbs, D. (2013). Small businesses and flood impacts: The case of the 2009 flood event in Cockermouth. *Journal of Flood Risk Management*, *7*(1), 42–53.
Weeg, C., Schwartz, H. A., Hill, S., Merchant, R. M., Arango, C., & Ungar, L. (2015). Using Twitter to measure public discussion of diseases: A case study. *JMIR Public Health and Surveillance*, 1(1), e6.
WHO Director-General's Report to Member States at the 76th World Health Assembly – 22 May 2023 (2023, May 22). https://www.who.int/director-general/speeches/detail/who-director-general-s-report-to-member-states-at-the-76th-world-health-assembly---22-may-2023
Williams, C. Y., Li, R. X., Luo, M. Y., & Eitutis, S. T. (2023). Exploring patient experiences and concerns in the online cochlear implant community: A cross-sectional study and validation of automated topic modelling. *Clinical Otolaryngology*, *48*(3), 442–450.
Wilton, R., Páez, A., & Scott, D. M. (2011). Why do you care what other people think? A qualitative investigation of social influence and telecommuting. *Transportation Research Part A: Policy and Practice*, *45*(4), 269–282.
Worldometers (2023). COVID-19 Coronavirus pandemic. https://www.worldometers.info/coronavirus/.Accessed 18th October 2023.
Xue, J., Chen, J., Chen, C., Zheng, C., Li, S., & Zhu, T. (2020). Public discourse and sentiment during the COVID 19 pandemic: Using latent Dirichlet allocation for topic modeling on Twitter. *PLOS ONE*, 15(9), e0239441.
Yin, H., Song, X., Yang, S., & Li, J. (2022). Sentiment analysis and topic modeling for COVID-19 vaccine discussions. *World Wide Web*, *25*(3), 1067–1083.
Yu, Y., & Wang, X. (2015). World Cup 2014 in the Twitter World: A big data analysis of sentiments in U.S. sports fans' tweets. *Computers in Human Behavior*, *48*, 392–400.
Yu, S., He, S., Cai, Z., Lee, I., Naseriparsa, M., & Xia, F. (2023). Exploring public sentiment during COVID-19: A cross country analysis. *IEEE Transactions on Computational Social Systems*, *10*(3), 1083–1094.
Zhang, T., Gerlowski, D., & Ács, Z. J. (2021). Working from home: Small business performance and the COVID-19 pandemic. *Small Business Economics*, *58*(2), 611–636.

Shaun Meric Menezes received his B.Tech. degree in Computer Science and Engineering from Manipal University, India, in 2021, and his M.Sc. degree in Digital Transformation and Innovation from the University of Ottawa, Canada, in 2024. He is currently working as a GIS/Database Specialist for the City of Whitehorse, Yukon, Canada. His current research interests include Artificial Intelligence and Natural Language Processing.

Ashok Kumar is a Ph.D. candidate at the Telfer School of Management, University of Ottawa, specializing in Digital Transformation and Innovation. He received his B.Tech. (Hons.) in Manufacturing Science and Engineering, and an M.Tech. in Industrial Engineering and Management (both in 2022). His current research interest explores the intersection of finance and natural language processing, where he leverages neural networks to analyze large-scale textual data.

Shantanu Dutta received a Ph.D. in Management (Finance) from Carleton University, Canada in 2006. He also received an MBA degree from Asian Institute of Technology, Bangkok, and Master of Engineering Management (MEM) and B.Tech. (Hons.) from Indian Institute of Technology, Kanpur. He is currently a professor at the Telfer School of Management, University of Ottawa, Canada. His areas of interest includes application of AI in finance, corporate governance, and mergers and acquisitions.