Mining Geotagged Tweets: Tracking Spatiotemporal Variation of Mental Health in Canada during COVID-19 Pandemic

by

Charlotte Zhuoran Pan

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Science

in

Geography

Waterloo, Ontario, Canada, 2023

Author’s Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

This thesis explores and analyzes the evolution of the pandemic as a stressor on mental health in Canada through monitoring the sentiment polarity dynamics, emotion trends, and changes in keywords being discussed on Twitter, spanning between January 2020 to December 2022. Leveraging the surging amount of geotagged social media data, this study deploys a combination of machine learning, geospatial mapping, and social sensing as a new approach to observe, quantify and evaluate the evolution of national-wide emotion trends and psychological status along the COVID-19 pandemic timeline in Canada, interpret the underlying key factors and events, and thus inform us on how to mentally "re-start" in the post-pandemic era. The proposed methods include social sensing, large-scale sentiment polarity detection, emotion classification, keyword analysis, and kernel density mapping. The dataset after processing is consisting of 430,399 geo-tagged tweets discussing pandemic subjects posted by Canadian users from January 1, 2020 to December 31, 2022.

The results of this study reveal that the overall sentiment and emotion composition was the most optimistic during the early half of the pandemic, from the early spring of 2020 to the summer of 2021, and turned to decline from then to the end of 2022, sending a warning signal in public mental well-being. Beneath this trend, several driving events emerged, ranging from the declaration of state of emergency in March 2020, the peak of vaccine hesitancy in November 2020, the release of new vaccine mandate in January 2022 to the Freedom Convey lasting from January 2022 to February 2022. The results also indicate that there is an observable geospatial disparity in the shifting patterns and the overall mental health levels between Montréal, a French-dominant region, and Vancouver, Calgary, Edmonton, Toronto, and Ottawa-Gatineau, which are English-dominant or bilingual regions. Also, along with a delayed period of peaks and bottoms in sentiment polarity, Toronto is displaying a slightly different mood than the other English-speaking cities. Last but not the least, we propose two action strategies, promoting education on the importance of vaccine behaviours and rebalancing the COVID-19 restrictions, for boosting public confidence regarding the pandemic and rebuilding psychological resilience in the current post-pandemic era.

As the first work tracking the long-term mental health of Canada as a country during the pandemic, this study evidences the conclusion that as the global economy starts to recover and the number of cases becomes gradually under control with the availability of the vaccine, the public psychological condition is not lifting as fast as the economy and the physical health in today's post-pandemic world.

Acknowledgements

Many thanks to Professor Jonathan Li, my supervisor, for inspiring me to step into the fantastic world of geospatial science and endowing me with the precious chance to change my life trajectory. I wish to express my sincere gratitude and appreciation for being so lucky to pursue my research and education under your guidance. Thank you for your consistent support along this wonderful two-year journey and your inspiration that will encourage me throughout my life. Meanwhile, I would like to show respect to my MSc Thesis Examining Committee members, Dr. Michael A. Chapman, Professor at the Department of Civil Engineering, Toronto Metropolitan University and Dr. Jue Wang, Assistant Professor at the Department of Geography, Geomatics and Environment, University of Toronto Mississauga. Thank you for providing such constructive comments and suggestions for enhancing my thesis.

Many thanks to Dr. Jing Yao, Senior Lecturer at the Urban Big Data Centre, University of Glasgow, UK, Dr. Xiaoxiang Zhang, Professor at the College of Hydrology and Water Resources, Hohai University, China, Dr. Joe Repka, Professor at the Department of Mathematics, University of Toronto, Xiao Su, my internship supervisor at Misumi Group Inc., Dr. Steven A. Roberts, Professor at the Department of Geography and Environmental Studies, Wilfrid Laurier University, Dr. Robert Mason at the Department of Philosophy, University of Toronto, Dr. Hilary Bergsieker, Professor at the Department of Psychology, University of Waterloo, and Jessica Trickey, Ph.D. candidate at the Department of Psychology, University of Waterloo. Thank you for being such amazing mentors, letting me see the sky above me, and illuminating my path forward.

Special thanks to Dr. Siqin Wang, Associate Lecturer at the University of Queensland, Australia and Dr. Sabrina L. Li, Assistant Professor at the School of Geography, University of Nottingham, UK for sharing such insightful feedback that inspired me to better display this work. Many thanks to my graduate student fellows, Yiqing Wu, Sarah N. Fatholahi, Hongjie He as well as my peers and friends in the Geospatial Intelligence and Mapping Lab at the University of Waterloo. Thank you all for the lakes and mountains we overcame together.

Even more thanks to my dear grandma, Zhou Rongfang. It is your unwavering love and support that let me become the best version of myself.

Dedication

I dedicate this work to all the kids who never gave up on their dreams.

Table of Contents

Author's Declaration ii

Abstract iii

Acknowledgements iv

Dedication v

List of Figures viii

List of Tables ix

List of Abbreviations x

Chapter 1 Introduction 1

1.1 Study Significance 1

1.2 Objectives and Research Questions 2

1.3 Thesis Structure 3

Chapter 2 Background and Related Studies 4

2.1 Background of Social Sensing 4

2.2 Social Media's Role and Impact on Public Health 5

2.3 Twitter and Geo-tagged Tweets 7

2.4 Methods for Sentiment Analysis 8

2.4.1 Polarity Detection in Sentiment Classification 10

2.4.2 Cross-Domain Classification in Sentiment Classification 10

2.5 Methods for Spatial Analysis 11

2.6 Chapter Summary 12

Chapter 3 Methodology 14

3.1 Study Area and Datasets 14

3.1.1 Study Area 14

3.1.2 Data Extraction 15
3.1.3 Raw Dataset 16
3.2 Workflow of Proposed Methodology 18
3.3 Data Validation 19
3.4 Data Framing 20
3.5 Methods of Sentiment Analysis 21
3.5.1 Sentiment Polarity Detection 21
3.5.2 Cross-Domain Emotion Classification 22
3.6 Methods of Spatial Analysis 22
3.7 Chapter Summary 23
Chapter 4 Results and Discussion 24
4.1 Sentiment Polarity Detection: A Longitudinal Study of Mental Health 24
4.2 Cross-Domain Emotion Classification: A Longitudinal Study of Mental Health 27
4.3 Word Cloud Mapping: A Longitudinal Study of Keywords 28
4.4 Kernel Density Mapping: A Spatiotemporal Study of Mental Health 33
4.5 General Discussion 38
4.6 Chapter Summary 40
Chapter 5 Conclusions and Recommendations 42
5.1 Conclusions 42
5.2 Public Health Implications 43
5.3 Limitations and Recommendations for Future Research 44
References 46
Appendix A Supplemental Material 53

List of Figures

Figure 2.1 Framework of COVID-19 social media studies on public health and epidemic (Source: Tsao et al., 2021) 5

Figure 2.2 Sentiment analysis: An overview of tasks and focuses (Source: Wankhade et al., 2022) 9

Figure 2.3 A treemap visualizing the number of words associated with each emotion type in NRC Emotion Lexicon (Source: Mohammad & Turney, 2013).......... 11

Figure 3.1 2021 Canada population by city (bubble size corresponds to population divided by 5,000) 15

Figure 3.2 Spatial distribution of the raw data points.......... 16

Figure 3.3 Flowchart of the proposed methodology.......... 18

Figure 3.4 Number of new GeoTweets discussing COVID-19 by month.......... 20

Figure 4.1 Mean sentiment score across COVID-19 stages in Canada 24

Figure 4.2 Monthly and daily mean VADER sentiment score in Canada 26

Figure 4.3 Overall emotion classification of all tweets from January 2020 to December 2022 27

Figure 4.4 Monthly and daily evolution of emotion in Canada.......... 28

Figure 4.5 Trending keywords for tweets with positive sentiment (framed in red) and tweets with negative sentiment (framed in blue) over seven stages 33

Figure 4.6 Kernel density ratio heatmap of positive and negative feelings across six major CMAs ... 34

Figure 4.7 Mean sentiment score by stage in six major CMAs.......... 35

List of Tables

Table 2.1 Social media data mining regarding the COVID-19 pandemic: An overview of platforms .. 7

Table 2.2 Representative sentiment analysis literature in crowdsourced data mining (Source: Niu & Silva, 2020) 7

Table 3.1 Keywords and hashtags defined in searching query........ 16

Table 3.2 Description of the raw dataset........ 17

Table 3.3 Number of observations for each COVID -19 stage in Canada 21

Table 4.1 Mean sentiment score across COVID-19 stages in Canada 25

Table 4.2 Mean VADER score and NRCLex emotion percentages by Stage in Vancouver 36

Table 4.3 Mean VADER score and NRCLex emotion percentages by Stage in Edmonton 36

Table 4.4 Mean VADER score and NRCLex emotion percentages by Stage in Calgary 37

Table 4.5 Mean VADER score and NRCLex emotion percentages by Stage in Toronto 37

Table 4.6 Mean VADER score and NRCLex emotion percentages by Stage in Ottawa-Gatineau...... 38

Table 4.7 Mean VADER score and NRCLex emotion percentages by Stage in Montréal 38

List of Abbreviations

ANEW	Affective Norms for English Words
ANOVA	Analysis of Variance
API	Application Programming Interface
CI	Confidence Interval
CMA	Census Metropolitan Area
COVID-19	Coronavirus disease, previously known as “2019 novel coronavirus”
GTA	Greater Toronto Area
HSD	Honestly Significant Difference
ICT	Information and Communications Technology
IoT	Internet of Things
LIWC	Linguistic Inquiry and Word Count
MVRD	Metro Vancouver Regional District, or simply Metro Vancouver
NLP	Natural Language Processing
NRCLex	National Research Council Canada Emotional Lexicon
SA	Sentiment Analysis
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SEM	Standard Error of the Mean
SMS	Short Message Service
SVM	Support Vector Machine
VADER	Valence Aware Dictionary for sEntiment Reasoning
WHO	World Health Organization

Chapter 1
Introduction

1.1 Study Significance

Out of the many major aspects in which the coronavirus disease of 2019 (COVID-19) has challenged our lives, there is one change going on that is less visible yet overwhelming in each corner of the world, the mental health crisis (Sampogna et al., 2021). Among the general populace, high levels of distress have been detected since the global outbreak of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2, the academic name of the virus causing COVID-19). Additionally, a particular form of COVID-19-related health anxiety disorder has been identified, which has been linked to a longer-term increased risk of utilizing mental health services. Social isolation, symptoms of depression and anxiety, and sleep problems have all been reported to become more common in young people and teenagers (Sampogna et al., 2021). All of these phenomena indicate that we are living in an era of profound transition, and it is important for scholars to pay close attention to the changes in mental health. However, monitoring mental health dynamics on a national scale requires collecting voices from massive groups of people along each stage of events. While the go-to method in current mental health research, survey-based assessments, can offer us less flexibility and promptness in achieving that (Wang et al., 2022), the social sensing method, equipped with the surging amount of geo-tagged social media data, is of great potential in helping researchers detect and quantify the national trends of emotions and the psychological status, understand the core events under the dynamics and even more, answer the key question: what is causing the mental health disparity from places to places and how can we deploy targeted strategy to the m ost needed places to sustain a psychological resilient society.

As evidenced by recent studies (Wang et al., 2022; Niu & Silva, 2020; Yang, 2021), social media with geo-tag function (e.g., geo-tagged Tweets, Sina Weibo, Instagram) and their derivative application programming interface (API) products make it possible for researchers to track the signals of mental status both individually and at broader level aggregated by locations, and even more, it can enable the mental health assessment at a national scale and covering a longer timeline of the pandemic to reflect public reaction each time policy or vaccine availability changes. By analyzing the retrieved blog contents using sentiment analysis, topic modelling, keyword analysis and many other natural language processing (NLP) algorithms, researchers can portray a dynamic view of the public's attitudes and psychological well-being at various aggregation levels.

In Asia, Australia, the US, and Europe such initiatives on national-scale mental health evaluation have been made, but not in Canada (Guntuku et al., 2020; Leng et al., 2021; Valdez et al., 2020; Wang et al., 2022; Yin et al., 2020). Furthermore, despite the fact that leveraging the power of social media data has grown into a popular research trend in the field of COVID-19 mental health analysis, it is pertinent to note that the majority of the literature in the academic field primarily focuses on the early stages of the epidemic, creating a limited perspective on time. Additionally, just a few studies decided to include geographic information in their list of features, giving geo-tagged social media data a tremendous chance to add a fresh viewpoint to the field of traditional mental health research.

All three of the aforementioned factors highlight the urgent need to make use of the wealth of data available on social media to expand the scope of the mental health assessment to a wider span of time during the pandemic and particularly in Canada, where mental health issues in the COVID-19 pandemic are less studied. This thesis fills such a gap.

1.2 Objectives and Research Questions

This paper aims at exploring the current advantages and challenges of geotagged social media data mining methods in the mental health research context and their specific implementation in Canada. To achieve this research target, we deploy a combination of machine learning and sentiment analysis algorithms as well as geospatial mapping method, navigating the following four research questions.

1. How does the social media data inform us about mental health dynamics along the Canadian pandemic timeline?

2. What are the critical events driving the trends in mental health? For example, is there a significant correlation between the trends and the publishing of a new pandemic policy? Or does it correlate with vaccine availability, COVID-19 cases or hospitalizations?

3. Is there an observable geospatial heterogeneity of mental health status during the same period across different metropolises in Canada?

4. What public health strategies and government response could be helpful to increase public optimism and confidence regarding the pandemic and rebuild psychological resilience in the current post-pandemic era?

1.3 Thesis Structure

This thesis is organized into five chapters.

Chapter 1 highlights the significance of the study while unfolding the research background, current challenges, and the research objectives of this study. A brief outline of the structure is also included.

Chapter 2 presents the research directions, data sources, and analytical methods of related social sensing studies in the field of public health and mental health research. The chapter introduces the state-of-the-art sentiment analysis tools and explains how social sensing-based studies on mental health incorporate the spatial component into their experiments through bivariate kernel density estimation.

Chapter 3 describes how the framework for this thesis is developed to answer the key research questions about mental health signals during the COVID-19 pandemic in Twitter Canada, including how datasets are created, how data are validated, how data are structured, and how analysis procedures are designed.

Chapter 4 reports, interpret and analyzes the outcome of sentiment analysis, emotion classification, keyword analysis and kernel density heatmap. The possible association between the observed patterns in sentiment polarity and crucial pandemic events such as government recognized COVID-19 as emergency, vaccines became available in Canada, public’s vaccine hesitancy changed, as well as government imposed stricter pandemic policies are also explored.

Chapter 5 concludes the thesis with key contributions indicated and provides thorough insights and discussion regarding the direction of future work.

Chapter 2
Background and Related Studies

In order to examine the relevant studies that have been done on the subject of social sensing-based mental health research, an overview of the literature is given in this chapter. The idea of social sensing is presented in Section 2.1, along with how it specifically applies to social media data mining in todays' academic research. The current research directions of mental health studies powered by social sensing as well as the factors influencing the selection of the optimal social media platform are subsequently covered in Section 2.2. Derived from those factors, Section 2.3 explains why Twitter and Geo-tagged Tweets suit best the data need for mental health studies in social media. Next, Section 2.4 provides a concise summary of the sentiment analysis methods along with the effectiveness of these methods in related works, and Section 2.5 sorts out the spatial analysis methods deployed in literature such as kernel density estimation, which enables researchers to evaluate the geographical differences in social media output. Section 2.6 concludes this chapter with key takeouts.

2.1 Background of Social Sensing

As a compounding effort of multiple disciplines— sociology, spatial science, sensor networks, social networks, cognition, data mining, information theory, linguistics, machine learning, behavioral economics and many others, social sensing is a relatively new term to both the academic and the industry field (Wang et al., 2019). Crowdsourcing data has been pumped to an unprecedented scale and granularity with the 21st-century technological revolution in Information and Communications Technology (ICT) and Web 2.0 (Niu & Silva, 2020; Yang, 2021). Such individual-level big geospatial data combined with related analysis techniques together form the methodology of social sensing (Yang, 2021; Wang et al., 2019).

To classify its nature, we are going to unravel the term social sensing in a frame of three layers, what are the "sensors" being used here, what kind of formats it outputs, and what new research direction it opens for us. As opposed to remote sensing which aims at helping us capture the physical aspects of the world, social sensing is proposed to help us detect the socioeconomic aspects of the world, and how human activity interacts with the environment. In social sensing, the mediums vary a lot, ranging from mobile phones and social media platforms to the Internet of Things (IoT) devices. However, people are playing a more critical role in the process of social sensing than these mediums and therefore, the sensor in social sensing data is realized by each individual in the sensor network (Wang et al., 2019; Liu et al., 2015). Besides, the available types of outputs are no longer limited to satellite imagery but include texts, sounds, videos, trajectories, and many others (Liu et al., 2015). Such flexibility in the outcome format and focus on human-environment interaction makes social sensing an ideal methodology for researchers to model social dynamics which are constrained by the empirical reality of human social systems (Galesic et al., 2021).

2.2 Social Media's Role and Impact on Public Health

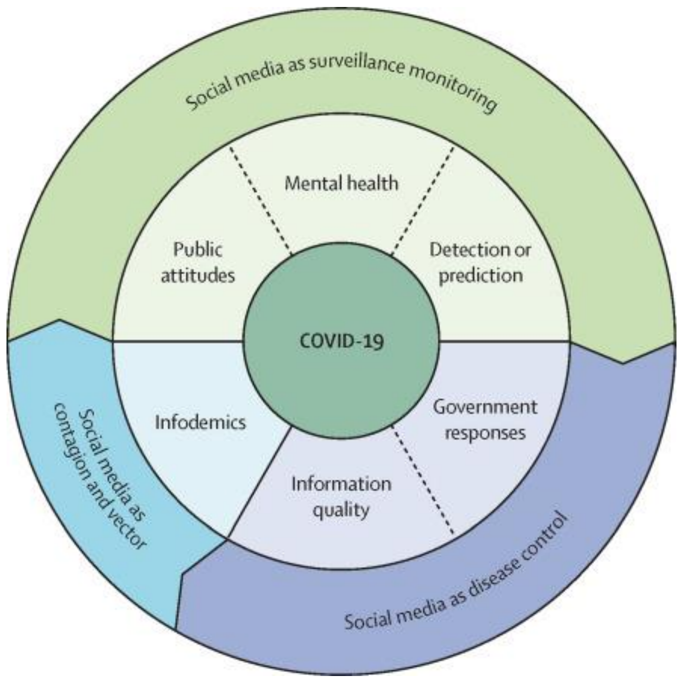


Figure 2.1 Framework of COVID-19 social media studies on public health and epidemic (Source: Tsao et al., 2021)

Figure 2.1 lists the major research themes of COVID-19 studies powered by social media in the public health domain classified by Tsao et al. (2021). As early as before the pandemic, Niu and Silva (2020) proposed social sensing big data and social media data as one of the core research interests and expertise in human geography domain while doing a systematic review of crowdsourced data mining for urban activity between 2013 and 2019. Later when the coronavirus has started to spread, a group of scholars, Tsao et al. (2021), noticed the increasing importance of social media in public health study regarding the COVID-19 pandemic and they summarized its application into three overarching umbrellas and six subordinate research themes, which are social media as surveillance monitoring (surveying public attitudes, assessing mental health, detecting or predicting COVID-19 cases), social media as contagion and vector (identifying infodemics), and social media as disease control (analyzing government responses to the pandemic, and evaluating quality of health information in prevention education videos). Likewise, another group of researchers, Huang et al. (2022b), categorized the social media data mining studies in the COVID-19 context into six major directions in order to systematically review their progress and limitations. Besides the previously concluded themes ranging from early warning and detection, communication and information conveying, public attitudes and emotions, to infodemic and misinformation, they also identified two new research areas, human mobility monitoring, and hatred and violence.

Table 2.1 lists the social media sites that have previously served as the primary data sources for COVID-19 public health studies. The majority of social media sites offer sample data sets or application programming interfaces (API) to access feeds from social media in the past or in the current moment. In the mental health, emotion, and public attitudes analysis subdomain of coronavirus literature, the two blogging platforms, Twitter and Sina Weibo, were the two most common social network sites being mined, according to the earlier scoping reviews (Niu & Silva, 2020; Huang et al., 2022b; Tsao et al., 2021). For sentiment analysis studies in general, Twitter data was more often used than data from other social media, or non-social-media crowdsourced databases which include products from points of interest (POIs) projects and collaborative websites (Niu & Silva, 2020; Tsao et al., 2021). The popularity of Twitter data in this field is further demonstrated by Table 2.2, which provides a summary of the representative literature of sentiment analysis in crowdsourced data mining from 2013 to 2019 as compiled by Niu & Silva (2020).

Table 2.1 Social media data mining regarding the COVID-19 pandemic: An overview of platforms

Category	Platform	Public Tools	Public Datasets if No Public Tools	Providing Geolocation	Providing Text	Geographic Coverage
Social networks	Twitter	Twitter Academic API (recommended), Twitter API	/	Yes	Yes	Global
	Sina Weibo	Not Publicly Available	Weibo COVID dataset (Leng et al., 2020)	No	Yes	Global; Posts mostly in China
Media sharing network	YouTube	YouTube Data API	/	No	No	Global
	Instagram	Instaloader, Instagram Hashtag search API	/	Yes	No	Global
Discussion forums	Reddit	PushShift API	/	No	Yes	Global
	Quora	Not Publicly Available	Not Publicly Available	/	/	Global

Table 2.2 Representative sentiment analysis literature in crowdsourced data mining (Source: Niu & Silva, 2020)

Research topic	Methods	Dataset	Case study	References
Happiness	Maximum entropy classifier	Twitter	London, UK	Quercia et al. (2012)
	Language Assessment by Mechanical Turk word list	Twitter	US	Mitchell et al. (2013)
	Language Assessment by Mechanical Turk word list	Twitter	US	Frank et al. (2013)
	Multinomial Naïve Bayes classifier	Twitter	New York City, US	Li et al. (2017)
	Unigram-based sentiment analysis	Twitter	London, UK	Guo et al. (2016)
Smart governance	AFINN dictionary	Twitter	US	Hollander and Hartt (2018)

Note. All of the listed representative works took Twitter as their only data source.

2.3 Twitter and Geo-tagged Tweets

There are a number of factors that set Twitter apart from other social media options. The data scale comes first. As of December 2022, there were more than 368 million active Twitter users worldwide (Dixon, 2022). Researchers are guaranteed both the scale of the population and the effectiveness of data collection thanks to this sizable user base and its readily available application programming interface (API) tools, Twitter API for public use and Twitter Academic API for qualified academic applicants.

Meanwhile, unlike some other social media platforms that can only be tracked for a short period of time, Twitter enables researchers to track tweets posted just a moment ago as well as all the way back to 2006 since Twitter first launched. This in turn fertilized a wide range of research in the mental health domain, either investigating emotion (e.g., level of happiness, anxiety, stress, depression), public's attitudes towards COVID-19 policies, (e.g., social distancing and school closure), reaction towards politicians, the public's awareness of COVID-19 related events (e.g., protests against lockdown, vaccination, and university reopening) or smart governance (Huang et al., 2022b).

In addition to the large size of data and flexible time range, Twitter is one of the few social networking sites that allows researchers to track both geolocation information and text content of a post in "Geo-tagged Tweets". "Geo-tagged Tweets", sometimes also referred to as " GeoTweets", are tweets that have location tags when they are published. This is especially critical to this study since we are interested in the spatial variation of the general public's mental health status in Canada.

2.4 Methods for Sentiment Analysis

An active area of research in the field of machine learning and natural language processing is sentiment analysis (SA), which examines people's views, sentiments, inspections, attitudes, and emotions through the computational interpretation of subjectivity in text (Hutto & Gilbert, 2014). Opinion mining (OM) is another name for it. The most common application of this technique is to extract opinionated information from a mass of text and determine the polarity of its sentiment.

As shown in Figure 2.2, sentiment analysis, the root node in green, is not a single task but a big umbrella that embraces various tasks. These five distinct tasks, represented as second-level nodes in blue, range from subjectivity classification, sentiment classification, opinion spam detection, implicit language detection to aspect extraction (Birjali et al., 2021; Wankhade et al., 2022). Besides, there are three different focuses identified under the sentiment classification task, namely polarity determination, cross-language classification and cross-domain classification. They are portrayed as red third-level nodes in Figure 2.2. In this thesis, we will focus on the two specific tasks in the sentiment classification domain, polarity determination and cross-domain classification. Innovative representatives of both categories are introduced in Sections 2.4.1 and 2.4.2, respectively.

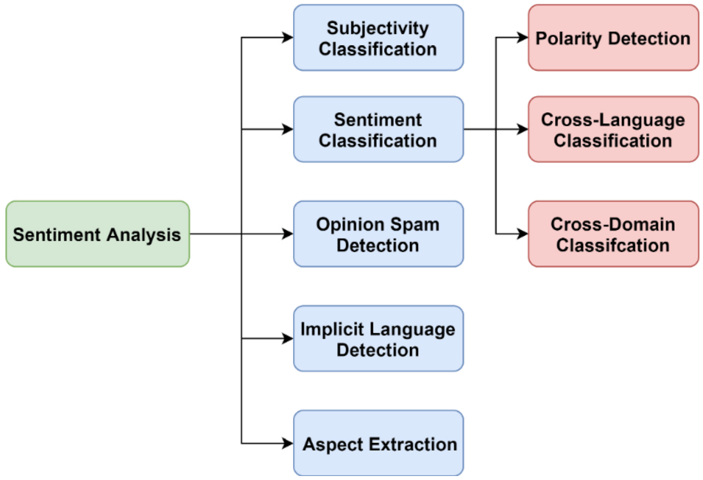


Figure 2.2 Sentiment analysis: An overview of tasks and focuses (Source: Wankhade et al., 2022)

Evidenced by recent reviews (Li & Wang, 2021; Umair et al., 2021), sentiment analysis has emerged as one of the most prevalent methods for evaluating mental health among various existing techniques, particularly when it comes to the task of interpreting mental well-being signals from social media data. Meanwhile, research and studies have demonstrated a strong correlation between the results of social media sentiment analysis and the ground truth psychological well-being of actual communities from a range of angles (Ajantha-Devi & Nayyar, 2021; Jaidka et al., 2020; Quercia et al., 2012 & 2021; Yin et al., 2020). Self-reporting has been found to be fairly reliable in evaluating one's well-being and naturally accords with the sentiment expressed on social media, so there is a solid basis for using social media content to measure people's mental well-being (Quercia et al., 2012). However, how accurately can such content reflect the mental health of the entire physical community, which includes people who seldom use social media? This issue was first brought up by Quercia et al. (2012), who conducted a comparison study between sentiment expressed in tweets and community socio-economic well-being matched by census communities in London, UK. They discovered a statistically significant association between the two, which led them to suggest that monitoring tweets is an effective way of tracking "gross community happiness," or the subjective well-being level of the entire community. Likewise, Jaidka et al. (2020) revealed that text analysis of social media produced robust and accurate indicators of regional well-being across the United States when compared to the gold-standard survey measures, Gallup-Sharecare Well-Being Index. In their case study, they combined 1.73 million Gallup survey responses and over a billion geolocated tweets, both from 2009 to 2015, to obtain county-level measures of life satisfaction, happiness, worry, and sadness. After post-stratifying the Gallup and Twitter samples to represent census demographics in age, gender, education, and income, they then tested and proved the validity of the Twitter outcome based on the Gallup survey results. As evidenced above, we conclude that researchers can employ large-scale sentiment analysis as a dependable sensor of community well-being and public mental health.

2.4.1 Polarity Detection in Sentiment Classification

One of the most commonly used polarity determination techniques in twitter mental health research is the Valence Aware Dictionary for sEntiment Reasoning (VADER) model (Umair et al., 2021; Wang et al., 2022). It is a machine-learning lexicon-based model that measures sentiment in a polar way as it estimates a single index of sentiment compound score ranging between two extremes, −1 (extremely negative) and +1 (extremely positive) (Hutto and Gilbert, 2014; Wang et al., 2022). Besides VADER, there are some other sentimental analysis models that can serve as benchmarks, such as Linguistic Inquiry and Word Count (LIWC), Affective Norms for English Words (ANEW), the General Inquirer, SentiWordNet, Naive Bayes-based machine learning models, Maximum Entropy, and Support Vector Machine (SVM) models. In the area of social media, the VADER lexicon performs remarkably well. With respect to the ground truth sentiment intensity of each tweet (aggregated group mean from 20 human raters), Hutto and Gilbert (2014) found that VADER (F1 Classification Accuracy = 96%) outperforms individual human raters (F1 Classification Accuracy = 84%), and generalizes more favourably across non-social media contexts than any of the previously mentioned benchmarks.

2.4.2 Cross-Domain Classification in Sentiment Classification

As a sub domain of sentiment analysis, cross-domain classification models are classifiers trained to predict the sentiment of a target emotion domain (Wankhade et al., 2022). A popular strategy to do this task is to extract the domain invariant features and their distribution. They can not only determine the sentiment of messages such as tweets and Short Message Service (SMS) in message-level task but also discover the multiple sentiment of a term within a message in term-level task (Mohammad & Turney, 2013).

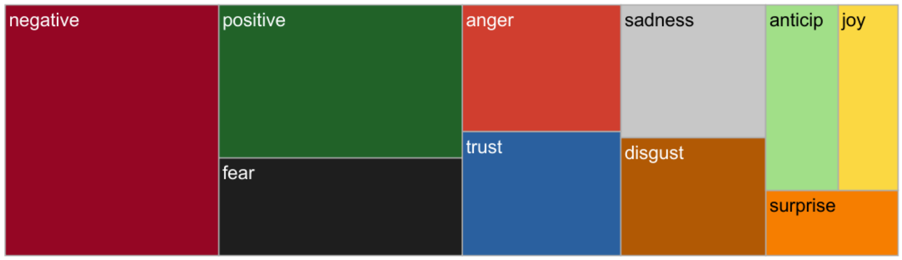


Figure 2.3 A treemap visualizing the number of words associated with each emotion type in NRC Emotion Lexicon (Source: Mohammad & Turney, 2013)

National Research Council Canada Lexicon (NRC Emotion Lexicon, a.k.a. NRCLex) is such a cross-domain emotion classifier that was developed on SVM and machine learning by researchers in National Research Council Canada (NRC) and Massachusetts Institute of Technology (MIT) that predicts the sentiments of a given text by documenting a list of English words and their associations with ten affects, comprised of eight basic emotions and two general attitudes. The two general attitudes are namely positive and negative. They link with the largest proportion of the total 27, 000 words in the NRC Emotion Lexicon among the ten categories of affections (see Figure 2.3). Eight basic emotions can be further grouped into four subtle pairs of bipolar emotions, consisting of joy (feeling happy) versus sadness (feeling sad); anger (feeling angry) versus fear (feeling of being afraid); trust (stronger admiration and weaker acceptance) versus disgust (feeling something is wrong or nasty); and surprise (being unprepared for something) versus anticipation (looking forward positively to something) (Mohammad & Turney, 2013).

2.5 Methods for Spatial Analysis

Some literatures were able to combine spatial analysis into their longitudinal studies of mental health variation, and a prevalent strategy for this is kernel density mapping, a geospatial mapping technique to visualize the density of geographic data (Hu et al., 2021; Krisp et al., 2009; Wang et al., 2022). To define what is kernel density mapping, we need to first meet the concept of kernel density estimation. Kernel density estimation (KDE) was first suggested by Rosenblatt (1969) to estimate the density of Y conditional on X = x where X is univariate and random. And it has nowadays become a classical non-parametric method of studying the density form of one or multiple random variable(s) (Chen, 2017;

Węglarczyk, 2018). The basic idea behind KDE is to estimate the probability density function by placing a kernel (a smooth function) at each data point and then summing up these kernels to obtain an overall estimate of the density (Chen, 2017). In one-dimensional implementation, the purpose of KDE is to generate a continuous univariate density function; while in case of the two-dimensional implementation, the goal of KDE turns to fit a raster-form density surface, which is a bivariate probability density function of the x- and y-coordinates, multiplied by the number of geographical features (Krisp et al., 2009; Węglarczyk, 2018). Here when the x- and y-coordinates we hope to link the density with are geo-coordinates, we call this fitted density surface a kernel density map, because it has geographical inference and can serve as a spatial analysis tool for point or polygon geographical features (Krisp et al., 2009).

The choice of amount of smoothness, termed as bandwidth in general or research radius in two-dimensional cases, is a core procedure in generating kernel density maps. Bandwidth selection matters a lot because too many wiggles and trivial details are displayed when the bandwidth is too small, important features are smoothed out when the bandwidth is too large, and we can clearly see the underlying density only when the bandwidth is set at a fair trade-off level (Chen, 2017; Węglarczyk, 2018). Typical ways to determine the bandwidth are the Silverman’s rule of thumb, least square cross-validation, biased cross-validation, and plug-in method (Chen, 2017). The most popular one, Silverman's bandwidth, is robust in case of spatial data outliers, and can determine the appropriate search radius at various spatial scales (Wang et al., 2022). Additionally, a proper colour/grey-scales classification method must be chosen in order to properly emphasize the distinction between relatively high and low values. Common choices include Jenks natural breaks classification, quantile classification and geometric classification (Krisp et al., 2009).

Application-wise, kernel density map can be utilized to figure out the distribution of a population based on a sample (as a method of interpolation and simulation), identify modes or peaks in the distribution (i.e., hotspots for vehicle accidents and neighbourhood crimes), and spot outliers or anomalies in the data (i.e., leaks in gas and oil pipelines) (Węglarczyk, 2018).

2.6 Chapter Summary

The research directions, data sources, and analytical methods of related social sensing studies in the field of public health and mental health research are discussed in this chapter. After reviewing the literature, we are able to conclude that geotagged tweets are the most effective social sensing data source for the COVID-19 mental health research taking place in Canada. And it has been established that employing sentiment analysis techniques is one of the best ways to capture mental health dynamics in blog-like content, particularly in tweets. The state-of-the-art sentiment analysis tools are also introduced here, along with each one's benefits and drawbacks. This chapter furthermore explains how social sensing-based studies on mental health incorporate the spatial component into their experiments through bivariate kernel density estimation, which is a prevalent geospatial mapping approach.

Chapter 3
Methodology

This chapter details the proposed methodology being utilized in this thesis. Section 3.1 outlines the geographical scope and introduces the datasets used in this study. Section 3.2 outlines the workflow of the methodology from the data gathering step to the conclusion-drawing step. Sections 3.3 and 3.4 walk through the data validation and data framing step in detail. The specific sentimental analysis algorithms we employ in this study are described in Section 3.5. Then the geospatial analysis methods are covered in Section 3.6. A brief summary of this chapter is provided in Section 3.7.

3.1 Study Area and Datasets

3.1.1 Study Area

In order to study the shifts in Canadian mental health status, this thesis choose Canada as the study region. According to the 2021 Census data ("Census of Population," 2023), Canada has a total of 36,991,981 population spreading among the ten provinces and three territories. There are 35 cities with a total population of at least 100,000 of which 50,000 or more must live in the core ("Canada: Metropolitan Area Population 2022 | Statista," 2023). Such a city is defined as a Census Metropolitan Area (CMA) in Census data. Out of the 35 CMAs, there are six super cities with over 1 million population, namely Toronto, Montréal, Vancouver, Calgary, Edmonton, and Ottawa - Gatineau ("Canada: Metropolitan Area Population 2022 | Statista," 2023). Canada's largest metropolitan area is Toronto, Ontario. In 2022, over 6.6 million people were living in the Greater Toronto Area (GTA). Greater Montreal, in Quebec, followed with about 4.4 million inhabitants, while the Metro Vancouver Regional District (MVRD, or simply Metro Vancouver), in British Columbia, counted 2.8 million people as of 2022. Almost 18 million people live in those six major metropolises alone, which is close to half of Canada's total population of 36 million. The high proportion of population in urban areas urges us to pay extra attention to these major cities in the subsequent geospatial analysis of mental health dynamics.

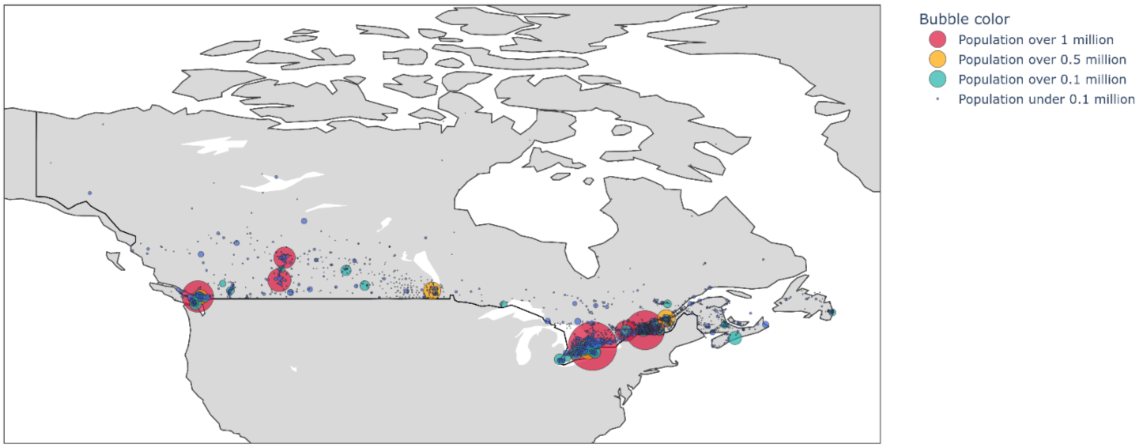


Figure 3.1 2021 Canada population by city (bubble size corresponds to population divided by 5,000)

3.1.2 Data Extraction

In this study, the geotagged tweets (GeoTweets) from Twitter in Canada during January 1, 2020 to December 31, 2022 were used as our destination for data mining. For improved population coverage and data representativeness, the Twitter Academic API and its full archive search endpoint Twitter API v2 were selected as the search engine after comparing with the alternatives. Twitter Academic API has a thorough application mechanism that evaluates what researchers are going to collect and how they are going to use the data from Twitter to answer their research questions, and then gives qualified researchers a chance to retrieve all the tweets matching with their predefined queries since 2006 ("Twitter API for Academic Research | Products," n.d.). This contrasts with the standard Twitter API v1.1, which always returns a random 1% of all the matched tweets for any query to protect data privacy ("Search Tweets Version Comparison," n.d.).

In this query, key words such as "COVID-19*, coronavirus, corona, pandemic, epidemic, virus, vaccine" were identified as the searching keywords and their related hashtags as searching hashtags (the asterisk * is a wildcard that can be replaced by any character or phrase). "2020-01-01 00:00:00 " was set as the start time and "2023-01-01 00:00:00 " as the end time which in turn give us a full three-year query time scope during January 1, 2020 to December 31, 2022. Other query terms include "CA" (Canada) as the nation and "EN" as the language (English). Those keywords listed above are a union of the terms identified as core words indicating COVID-related discussion and used to retrieve the development of pandemic topics on Twitter by Atillo (2021), Huang et al. (2022a), Valdez et al. (2020), Wang et al. (2022), and Yin et al. (2020), separately. In order to control repeated content, this query also adds a "-is:retweet" command to exclude retweets (the hyphen - is a logical sign standing for negation).

Table 3.1 Keywords and hashtags defined in searching query

Keywords	Hashtags
COVID-19 OR coronavirus OR corona OR pandemic OR epidemic OR virus OR vaccine	#COVID OR #COVID-19 OR #COVID19 OR #coronavirus OR #pandemic OR #vaccine OR #vaccin OR #immunization OR #vax

3.1.3 Raw Dataset

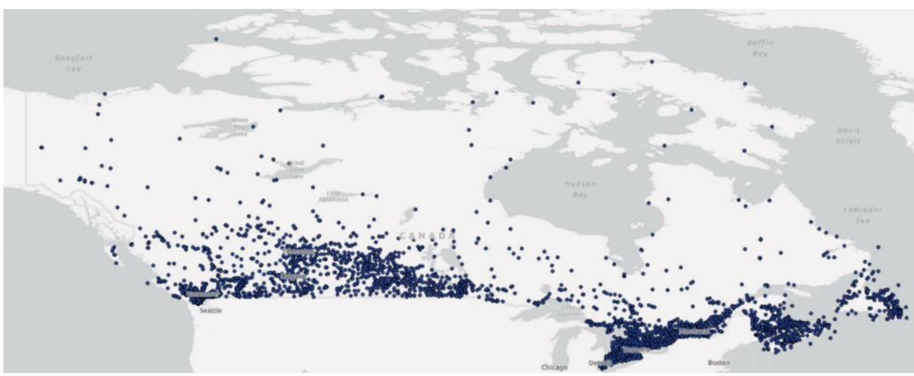


Figure 3.2 Spatial distribution of the raw data points

A total of 459,824 GeoTweets regarding coronavirus posted by 62,267 distinctive users were initially retrieved through the query defined above using full archive search within Twitter Academic API. The spatial distribution of the raw data is visualized in Figure 3.2. The geocoded location of data points spread across the ten provinces and three territories of Canada, however, they are concentrated notably more by the south border, especially within Southern Ontario by the Great Lake. This pattern matches the general distribution of Canada population portrayed in Figure 3.1.

For each GeoTweet, we gathered the following information listed in Table 3.2.

Table 3.2 Description of the raw dataset

Parameter	**Type**	**Description**
content	string	Text of the tweet in UTF-8 format.
create_at	date	The date and time of which this tweet is created.
author_id	string	The unique identifier of the user who posted this tweet.
username	string	The unique screen name that this user identifies themselves with.
profile_location	string	The location specified in the user's profile if the user provided one.
place.id	string	The unique identifier of the place tagged in the tweet.
place.name	string	The short name of this place, e.g., "Ottawa".
place.full_name	string	The detailed place name in a longer form, e.g., "Ottawa, Ontario".
place. country_code	string	The ISO Alpha-2 country code this place belongs to.
place.geo	string	Contains place details in GeoJSON format.
place.type	string	The particular type of location represented by this place information, such as a city name, or a point of interest.

There are two types of geographical tags in GeoTweets. Users can embed their location in a tweet by either tagging a precise coordinate or tagging a Twitter "place" object.

The first option, a precise coordinate, refers to a pair of longitude and latitude values in decimal degrees that designates a specific dot on Earth based on the World Geodetic System 1984 (WGS84) coordinate system. This option was cancelled from the Twitter iOS or Android app in June 2019, but it is still possible to use it with some third-party clients, through the API, and when using the in-app camera on Twitter for iOS and Android (Twitter, n.d.).

A Twitter "place", the second option, is a specific, named location with corresponding geo coordinates. It is a collection of four elements -- name, type, country code, and bounding box. By a

Twitter "place", users can choose between a polygon location (this way the bounding box will be made up of four WGS84 longitude and latitude coordinates in style [LONG, LAT, LONG, LAT]), which can be anything from a small neighborhood to a whole nation, or a point type location (the bounding box will be a pair of WGS84 longitude and latitude coordinates, e.g. [LONG, LAT]), like a movie theatre (Twitter, n.d.). Examples of both formats of location information in Twitter Academic API output are provided in Appendix A Section 1.

3.2 Workflow of Proposed Methodology

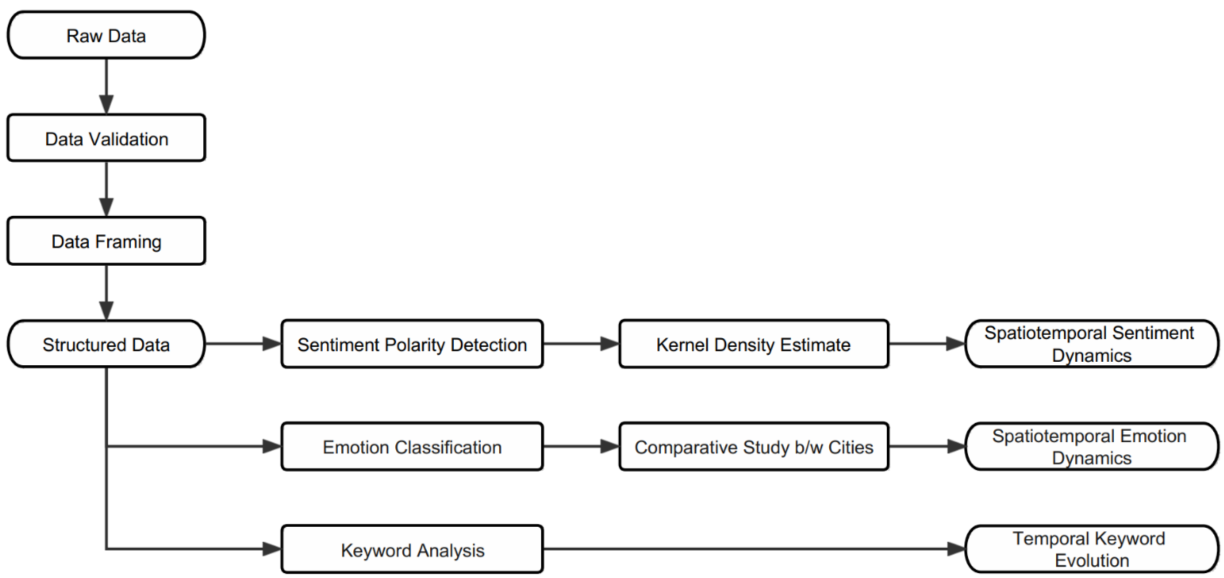


Figure 3.3 Flowchart of the proposed methodology

Note. Data inputs and outputs are denoted as ovals; Process steps are denoted as rectangles.

Figure 3.3 shows the flowchart that symbolizes the proposed methodology. The initial pre-processing is carried out prior to any actual analysis to ensure that the data fed into the sentiment analysis algorithms is of high quality and can accurately represent individual emotion. This involves filtering out content published by news organizations' accounts as well as removing emojis and links in the text. Next, the processed data are encoded into a structured dataset so that it can be organized by stage, month, and date. This step splits the three years' worth of tweet data into seven stages based on the peaks and troughs of coronavirus infection cases.

In order to assess the mental well-being dynamics from the twitter data, the first step is to run the VADER model to compute a sentimental compound score with a range of [-1, 1] for each tweet. The next step is to group the findings by day, month, and stage to investigate the longitudinal variation in the degree of optimism and pessimism among the public. Afterwards, the National Research Council Canada Lexicon (NRCLex) algorithm is applied to each post to assess the proportion of eight distinct emotions in each post. Once more, the outcome data are segmented by day, month, and stage to better visualize, compare, and interpret how the public's emotional responses change over time. Meanwhile, in order to examine the composition of emotions overall and by stage, the NRCLex model is also implemented to the text concatenating all tweets and tweets from each stage. The distribution of both positive and negative emotions in Canada's urban areas—which are the locations with the highest geo-tweet concentration—is then examined using a kernel density estimate tool.

3.3 Data Validation

Because the research goal is to track personal sentiments that can accurately reflect public mental health in our data, 29,425 posts from news organizations' accounts are excluded. This leaves 430,399 tweets remaining. Figure 3.4 depicts the count of new GeoTweets that mention the COVID pandemic grouped by month from January 2020 to December 2022 in Canada. It is important to note that in March 2020, discussions about the pandemic reached unprecedented heights. This calls for extra attention in considering how the data should be reframed into the timeline, which is dicusssed in the following section.

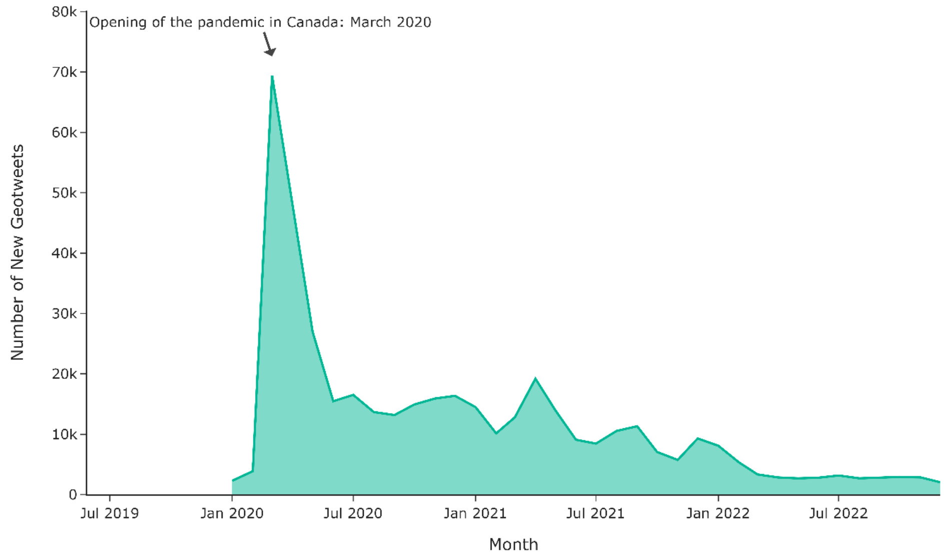


Figure 3.4 Number of new GeoTweets discussing COVID-19 by month

3.4 Data Framing

Table 3.3 shows the segmentation of the Canadian COVID-19 timeline. Both the starting and ending dates are inclusive. First, we divided the three-year span into seven periods into a framework of five infection waves and a possible sixth wave in the subsequent phase based on the *COVID-19 infection in the Canadian household population* published by Statistics Canada in April 2022 (Bushnik et al., 2022). The original duration of the first wave (2020-01-25 to 2020-07-17) defined by the Timeline was further divided into two phases, pre-pandemic and first wave, to reflect the fact that March 2020 was a unique turning point for Canada. On March 11, 2020, the World Health Organization (WHO) announced COVID-19 as a global pandemic (World Health Organization, 2020). During the following week of March 17, 2020, states of emergency were declared in most of the Canadian provinces and territories, marking the formal coming of the pandemic era in Canada (Luscombe & McClelland, 2020); On March 30, 2020, daily active cases national-wide peaked for the first infection wave, at a scale of 35,040 people (Bushnik et al., 2022). The starting point of the overall search timespan was stretched back to January 1, 2020 to enable a more comparable length of the pre-pandemic phase to the subsequent phases.

Table 3.3 Number of observations for each COVID -19 stage in Canada

Stages	**COVID-19 Timeline in Canada**	
	Duration	**No. of GeoTweets**
Stage 0, pre-pandemic	2020-01-01 to 2020-03-16	73,390
Stage 1, the first wave	2020-03-17 to 2020-07-17	130,969
Stage 2, the second wave	2020-07-18 to 2021-03-04	96,403
Stage 3, the third wave	2021-03-05 to 2021-07-22	52,447
Stage 4, the fourth wave	2021-07-23 to 2021-11-03	28,216
Stage 5, the fifth wave	2021-11-04 to 2022-02-28	24,569
Stage 6, post-pandemic	2022-03-01 to 2022-12-31	24,405

In the latter half of the pandemic, the fourth wave, the fifth wave, and the post-pandemic stages all share a similar number of observations (n) of about 25,000 GeoTweets after aggregation, whereas the four earlier stages have much larger cell sizes. The first wave, a crucial stage in our interest, has seen 130,969 GeoTweets alone, more than the number of tweets added together during the four stages from March 2021 to December 2022.

3.5 Methods of Sentiment Analysis

3.5.1 Sentiment Polarity Detection

To monitor the overall sentiment dynamics in Canada, the machine learning model Valence Aware Dictionary for sEntiment Reasoning (VADER) was applied to the GeoTweets we collected with a full length of 36 months. As previously mentioned in Section 2.4, it measures a sentiment compound score ranging between −1 (extremely negative) and +1 (extremely positive) for a piece of text input (Hutto & Gilbert, 2014).The compound score can then classify this piece of content into three genres: optimistic if the score is greater than 0.05, neutral if the score is between 0.05 and -0.05 or otherwise pessimistic if the score is below -0.05 (Hutto & Gilbert, 2014). Here we use utilize VADER to characterize the change in degree of optimism or pessimism across time and space.

3.5.2 Cross-Domain Emotion Classification

To detect change in more specific emotion rather than just optimism or pessimism, we decide to deploy National Research Council Canada Lexicon model (NRCLex). It is a machine-learning emotion classifier based on SVM and it predicts the associations of a given text with ten affects, comprised of eight basic emotions and two general attitudes (Mohammad & Turney, 2013). The two general attitudes are namely positive and negative. The eight basic emotions can be further structured into four subtle pairs of bipolar emotions, consist of joy (feeling happy) versus sadness (feeling sad); anger (feeling angry) versus fear (feeling of being afraid); trust (stronger admiration and weaker acceptance) versus disgust (feeling something is wrong or nasty); and surprise (being unprepared for something) versus anticipation (looking forward positively to something) (Mohammad & Turney, 2013). By monitoring the change in the share of four positive feelings versus the share of four negative feelings as well as the dynamic in the intensity of each emotion type alone, we can obtain further details about the evolution of public mindset as the pandemic going forward.

3.6 Methods of Spatial Analysis

Kernel density mapping is one of the most popular tools to compare and visualize differences between regions, as mentioned in Section 2.5, and here we are going to use it to speculate the spatial difference in sentiment polarity and identify hot spots. We noticed in Figure 3.2 that the number of GeoTweets in our sample is distributed very unevenly among different cities and inside the same city. Therefore, to control for the sample size, we are not going to apply the standard 2D implementation of kernel density estimation which constructs a density surface with single feature. Instead, we are going to calculate kernel density ratio and fit a relative density surface using two features, the sentiment polarity score and the sample size, as inputs.

In this approach, we first geocode the location of all GeoTweet points and select those points within the polygons of the six CMAs with the most population, Vancouver, Calgary, Edmonton, Toronto, Ottawa – Gatineau and Montréal. Next, we calculate the kernel density ratio for each raster cell to be the sum of sentiment scores divided by the number of GeoTweets within the search radius of this cell. The search radius or kernel's bandwidth controls how much smoothing is employed to the data; greater bandwidths produce smoother estimates but may sacrifice some data detail (Chen, 2017). To obtain the optimal bandwidth, we use a cross-validation bandwidth selection technique which targets multivariate data and was devised by Duong & Hazelton (2005). Our kernels are defined in a fixed way. In other words, our search radius and shape of the kernel are consistent across the map. In terms of the colour scheme methods, we take quantile (histogram equalize) classification to assign each hue in a blue-red colour band to each output value.

3.7 Chapter Summary

This chapter explains how the framework for this thesis is developed to answer the key research questions about mental health signals during the COVID-19 pandemic in Twitter Canada, including how datasets are created, how data are validated, how data are structured, and how analysis procedures are designed. To track the longitudinal changes in mental health status, this thesis adopts Valence Aware Dictionary for sEntiment Reasoning (VADER) model to estimate the sentiment polarity, and the National Research Council Canada Lexicon model (NRCLex) to account for specific emotions such as fear, trust and joy. The proposed methodology also includes a world could algorithm to study the top keywords in each stage of the pandemic. To explore the geospatial patterns in mental health status, kernel density ratio mapping based on the sentiment polarity results and sample size is incorporated in the methodology to assess the disparity among different metropolises of Canada.

Chapter 4
Results and Discussion

This chapter discusses the outcomes of the proposed methodology in detail. Section 4.1 displays, analyzes and interprets the change of optimistic and pessimistic sentiments shown in tweets across the time. Section 4.2 reports, visualizes, and assesses the estimated percentage of emotions by each pandemic stage. A keyword analysis via word cloud mapping is presented in Section 4.3, and a geospatial analysis via kernel density mapping is included in Section 4.4. Section 4.5 elaborates on the outcome relating to the findings of other relevant works and Section 4.6 concludes this chapter.

4.1 Sentiment Polarity Detection: A Longitudinal Study of Mental Health

The mean VADER sentiment scores by Stage and their corresponding 95% confidence intervals (CI) are presented in Table 4.1 and Figure 4.1 below. Overall, the scores resembled a vibrating mountain shape. During this three-year span, the stage average of sentiment score surged from the worst of 0.008 at the pre-pandemic stage, to as high as 0.082 at the first wave, dropped again slightly to 0.076, rose to the best in three year at 0.091 at the third wave, dropped again dramatically to 0.038, bumped up again to 0.044 at the fifth wave, and then collapsed to 0.015 at the most recent post-pandemic stage, foreshadowing a pessimistic tendency towards the future.

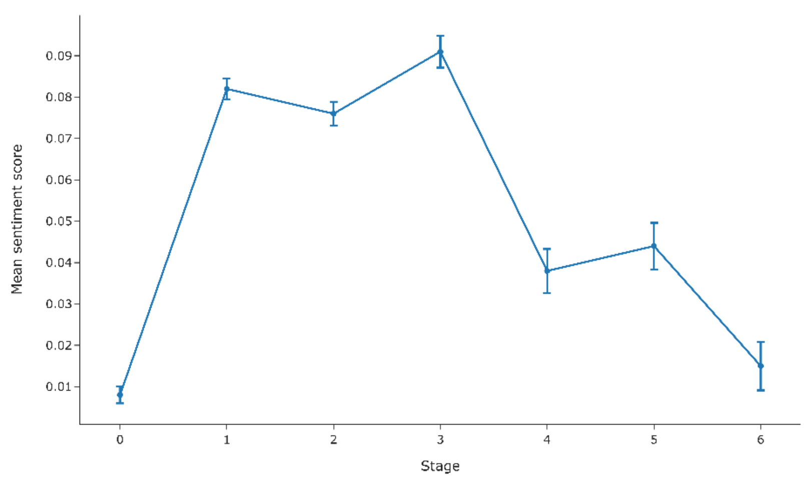


Figure 4.1 Mean sentiment score across COVID-19 stages in Canada

Note. Error bar denotes the upper bound and lower bound of the 95% confidence interval.

Table 4.1 Mean sentiment score across COVID-19 stages in Canada

Stages	**COVID-19 Timeline in Canada**		
	Duration	**Mean**	**95% Confidence Interval**
Stage 0, pre-pandemic	2020-01-01 to 2020-03-16	0.008	[0.006, 0.010]
Stage 1, the first wave	2020-03-17 to 2020-07-17	0.082	[0.080, 0.085]
Stage 2, the second wave	2020-07-18 to 2021-03-04	0.076	[0.074, 0.079]
Stage 3, the third wave	2021-03-05 to 2021-07-22	0.091	[0.087, 0.095]
Stage 4, the fourth wave	2021-07-23 to 2021-11-03	0.038	[0.033, 0.043]
Stage 5, the fifth wave	2021-11-04 to 2022-02-28	0.044	[0.038, 0.050]
Stage 6, post-pandemic	2022-03-01 to 2022-12-31	0.015	[0.009, 0.010]

To quantify the extent to which the Stage (0, 1, 2, 3, 4, 5, 6) interacts with the sentimental polarity, a repeated-measures Analysis of Variance (ANOVA) was conducted. A statistically significant main effect of Stage emerged, $F(6, 430392) = 366.97$, $p < .001$. Pairwise Tukey HSD post-doc tests revealed that, except Stage 0 and 6 (p = 0.362) and Stage 4 and 5 (p = 0.680), the mean sentiment scores are different between all pairs of stages (p = 0.019 for Stage 1 and 2, p = 0.003 for Stage 1 and 3, p < 0.001 for the rest pairs).

Among the six transition given by seven stages, only two stages, Stage 1 and Stage 3, show a statistically significant increase in sentiment score than the previous stage. During Stage 1 featured by the first wave happening in the second quarter of 2020, twitter users' confidence over pandemic topics embraced a sharp recovery, making a 935.75% hike to 0.082 (95% CI from 0.080 to 0.085) from the previous pre-pandemic stage 0.008 (95% CI from 0.006 to 0.010). One year later, the sentiment score climbed up again with a 18.89% increase from the 0.076 (95% CI from 0.074 to 0.079) in Stage 2, reaching the all-time highest at 0.091 (95% CI from 0.087 to 0.095) during Stage 3, the third wave in the summer of 2021, in which the COVID-19 vaccines became readily available in Canada (Bushnik et al., 2022). However, the positive emotion tide did not last long. Sentiment score fell off the mountain since Stage 3, stopped and then continued falling off to 0.015 (95% CI from 0.009 to 0.020), in the final stage. During the final post-pandemic stage which are the closet stage to today, the sentiment score shrank by a considerable amount, 66.72% of the previous stage, sending a warning signal for our current mental health status.

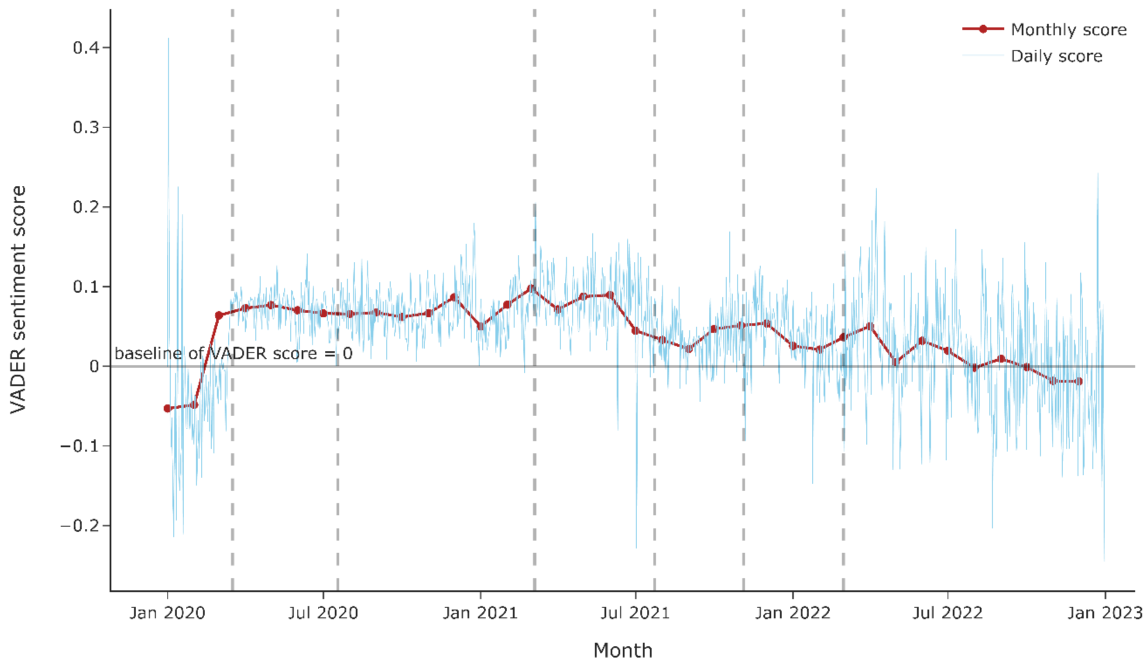


Figure 4.2 Monthly and daily mean VADER sentiment score in Canada

To dig deeper into the possible drivers behind the trend, we need to examine the variation of the monthly average sentiment score. Figure 4.2 illustrates the evolution of the monthly mean and daily mean VADER sentiment score of tweets throughout our research timespan from January 1, 2020 to December 31, 2022 in Canada. In the pre-pandemic phase starting from January 2020, the monthly sentiment was at the bottom. And then, as soon as most Canadian regions declared a state of emergency, the score rushed to the apex of three years in March 2020. Such an optimistic trend maintained all the way to the middle of 2021 and then gradually lose its rising power. A local minimum sentiment score appeared in February 2022, which overlapped the ending of Freedom Convoy (January 15 to February 14, 2022). Since then, the overall sentiment score kept sliding downwards as documented by our last recorded month, December 2022.

4.2 Cross-Domain Emotion Classification: A Longitudinal Study of Mental Health

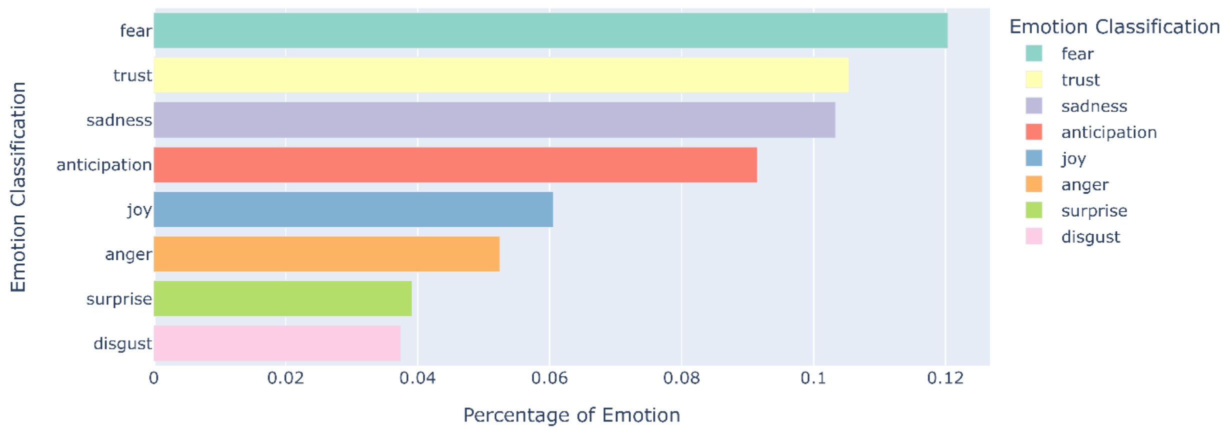


Figure 4.3 Overall emotion classification of all tweets from January 2020 to December 2022

Figure 4.3 displays the distribution of emotions in all English-language tweets about the pandemic sent within Canada over the course of three years. The percentage of a single emotion is calculated as the ratio of words that are linked with that emotion type from the joined text of all tweets in the dataset. According to the classification mechanism of NRCLex, the most prevalent emotions are fear (12.04%), trust (10.54%), sadness (10.33%), anticipation (9.15%), joy (6.05%), anger (5.24%), surprise (3.91%) and disgust (3.74%). Fear takes up the largest portion of the eight types of emotions in the classification of feelings over the span of three years and in almost every stage of the epidemic.

Figure 4.4 illustrates how the eight categories of emotions shifted in or faded out over the course of the pandemic. As time goes on, we witness that the two major pessimistic classes of emotion, sadness and fear, both exhibit a steady rise in their emotional share in general, while at the opposite end of the emotional spectrum, the two dominant optimistic feelings, trust and anticipation, are both in decline. Meanwhile, the two remaining negative emotion categories, anger and disgust, show a relatively flat trend along the timeline, so do the two remaining positive emotion types, joy and surprise.

From Stage 0 (pre-pandemic) to Stage 3 (the third wave), the changes in the share of emotions have been relatively stable, except for a sharp rise of fear and anger in February 2021 and a sharp drop of them two months later in May 2021. Later in the middle of Stage 4 (the fourth wave), the autumn of 2021, trust entered into an unprecedented leap, but anticipation moved in a reverse direction, dropping into a local minimum. Starting in Stage 5 (the fifth wave) which is marked by November 2021, the gap between the pessimism and optimism widens at a dramatic speed, with the percentage of fear and sadness added together flying streets ahead the share of trust and anticipation. When it comes to the current stage, Stage 6 (post-pandemic), fear still accounts for the largest proportion, occupying 14.25% of the length of all tweets posted during this period, followed by sadness (12.52%), trust (8.35%), anticipation (7.60%), anger (5.32%), joy (4.65%), disgust (4.13%), and surprise (3.61%). Overall, the four negative emotions are showing a much stronger power than the four positive emotions expressed by Twitter users.

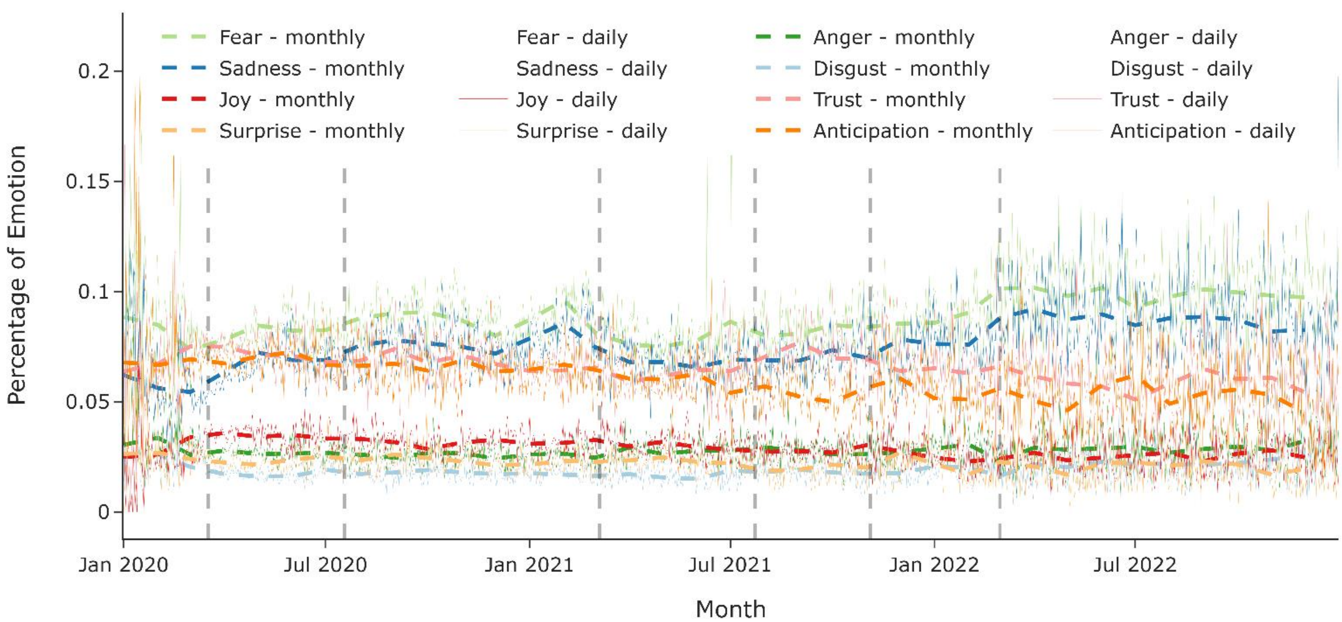
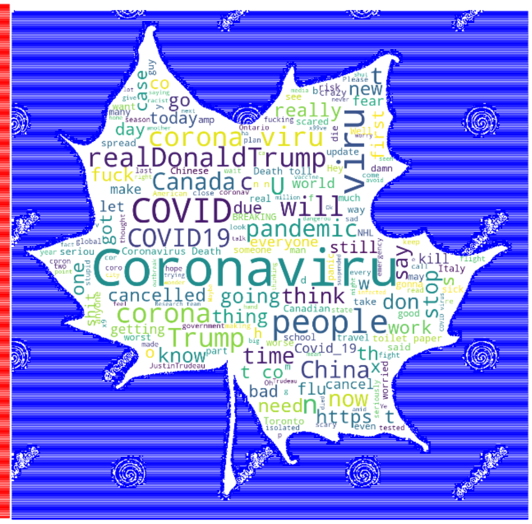


Figure 4.4 Monthly and daily evolution of emotion in Canada

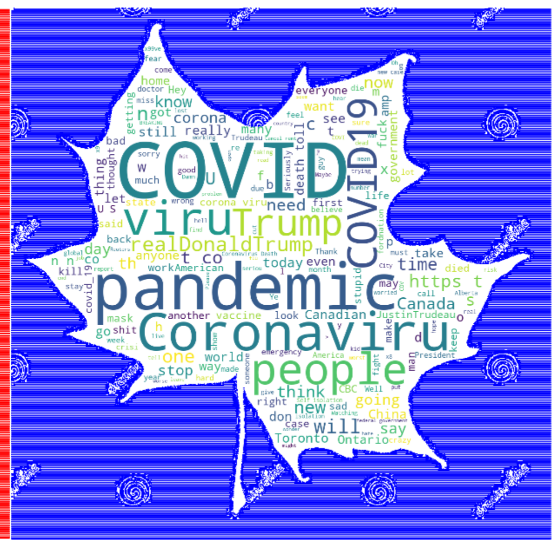
4.3 Word Cloud Mapping: A Longitudinal Study of Keywords

Figure 4.5 presents the world could of keywords appeared on Twitter under both upbeat and downbeat sentiments for each of the seven pandemic stages. To understand what people are concerned about when they talk about the pandemic, a world cloud mapping is performed on the text content of collected GeoTweets. In Figure 4.5, the most relevant keywords in the tweets recognized as optimistic mode (0.05 ≤ VADER score ≤ 1) are displayed on a red background, while those in the tweets detected as pessimistic mode (-1 ≤ VADER score ≤ -0.05) are framed in blue. Here the subtitles below each pair of graphs state the corresponding Stage, and the size of a word represents its popularity and the frequency with which it appears in tweets. Several COVID-related keywords appear from the beginning to the end under both sentiment contexts. While it is noticeable that in the early stages, people were referring to the COVID more using formal words, including “corona”, “virus”, “coronavirus” and “COVID19”; as the time passed by, people switched to more informal words such as “pandemic” and “COVID”. The word "vaccine" started popping up among the keywords of both positive and negative discussions in Stage 2 (2020-07-18 to 2021-03-04), and it quickly climbed to the top spot in Stages 3 (2021-03-05 to 2021-07-22) and Stage 4 (2021-07-23 to 2021-11-03) for both sentiments. The most often used words for posts with positive emotion were "thank," "people," "good," "great," "Canada," "today," "new," and "think." Some of these words, like "people," "Canada," "today," "new," and "think," also frequently appeared in negative scenarios. This could be explained by the fact that those phrases are spoken a lot in everyday conversation of Canadian. However, some of the terms were much more trending under the negative context than in the positive context. For example, “Trump” and “realDonaldTrump” were among the most frequent terms in negative discussion from Stage 0 (2020-01-01 to 2020-03-16) to Stage 2 (2020-07-18 to 2021-03-04). This could have been related to the 2020 US Presidential Election and Trump’s remarks about COVID during the beginning of the pandemic.



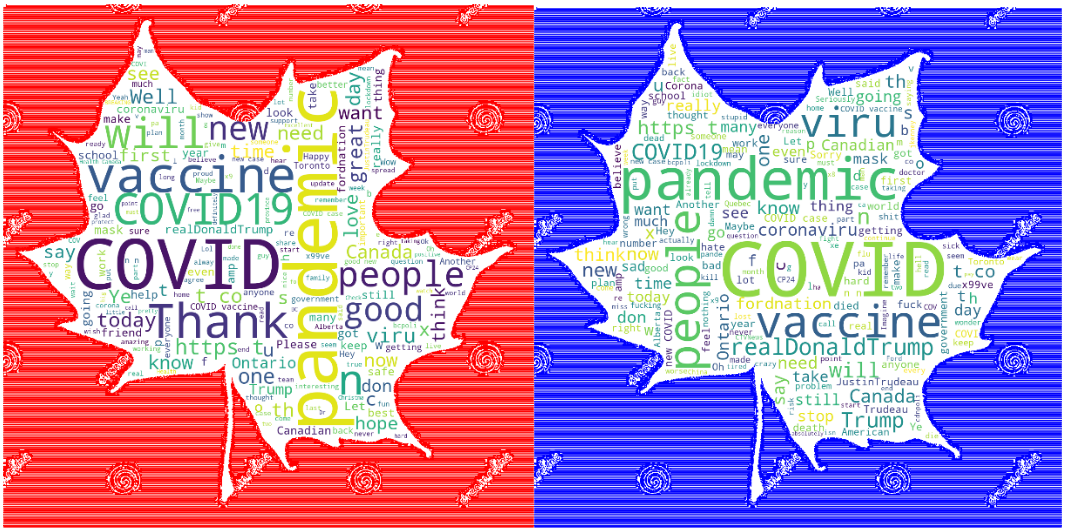
(a) Stage 0 optimistic

Stage 0 pessimistic



(b) Stage 1 optimistic

Stage 1 pessimistic

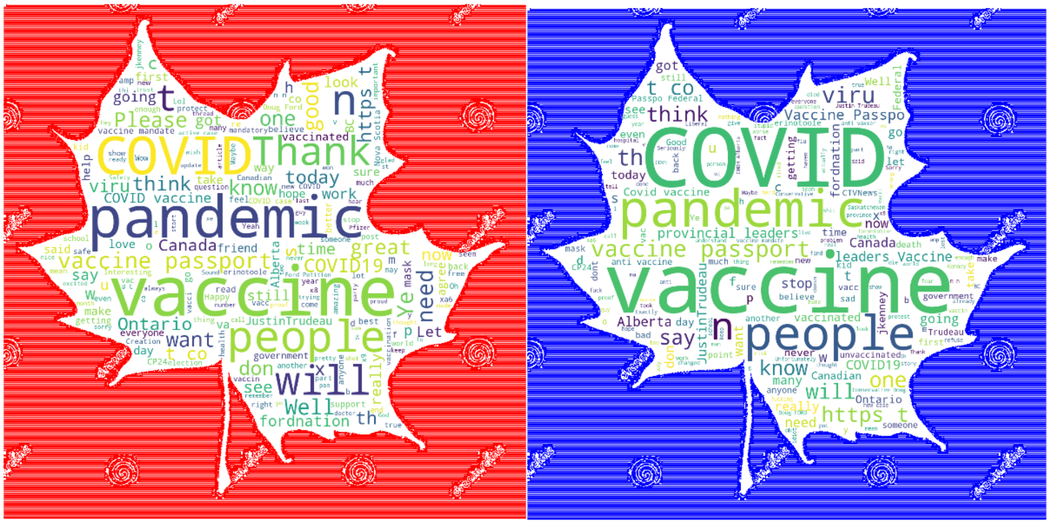


(c) Stage 2 optimistic

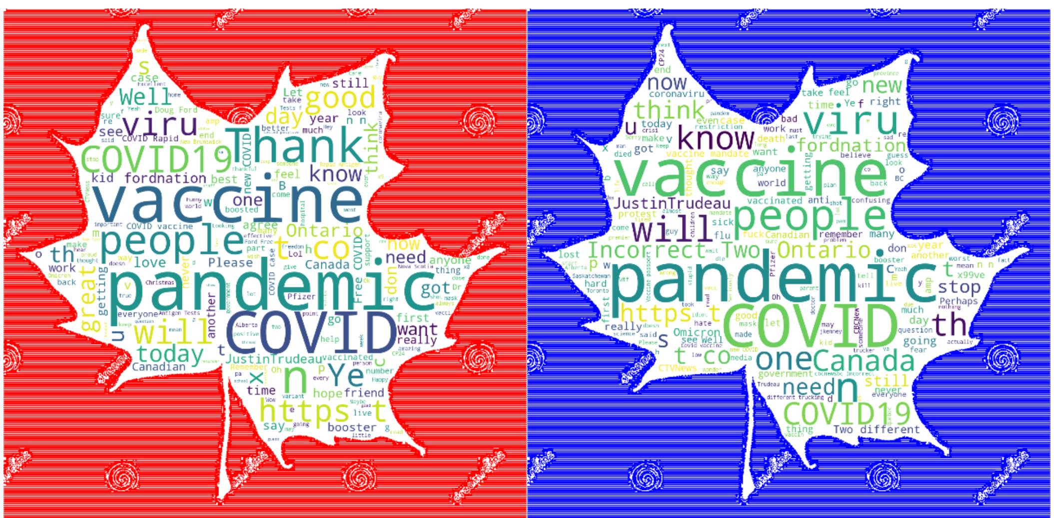
Stage 2 pessimistic

(d) Stage 3 optimistic

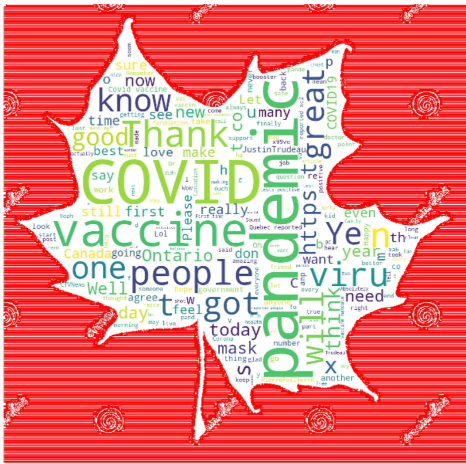
Stage 3 pessimistic

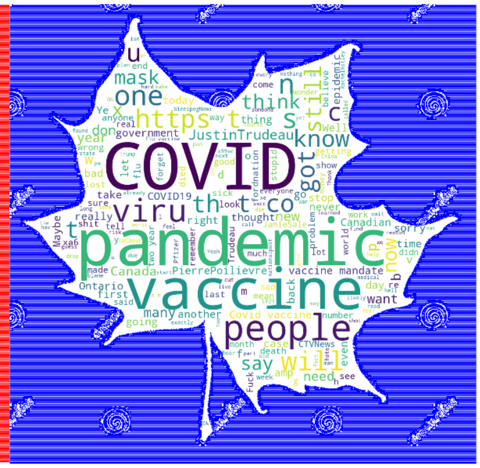


(e) Stage 4 optimistic Stage 4 pessimistic



(f) Stage 5 optimistic Stage 5 pessimistic





(g) Stage 6 optimistic

Stage 6 pessimistic

Figure 4.5 Trending keywords for tweets with positive sentiment (framed in red) and tweets with negative sentiment (framed in blue) over seven stages

4.4 Kernel Density Mapping: A Spatiotemporal Study of Mental Health

The plague, according to Albert Camus (1991), brings "impartial justice" to the city of Oran, although this may not be the case, especially when considering the plague has swept the country or the globe. Figure 4.6 is a kernel density ratio map that illustrates the spatial variation of sentiments among the six CMAs Vancouver, Edmonton, Calgary, Toronto, Ottawa-Gatineau, and Montréal, with optimistic hot spots (the concentration of positive sentiments) showing in vivid red and pessimistic cold spots (the concentration of negative sentiments) showing in dark blue. The distribution of hotspots and cold spots highlights the differences in the mental well-being signals among these metropolitan areas.

Observing the distribution of optimism, we found that all CMAs have at least one hotspot, while only Toronto and Ottawa - Gatineau display a hotspot pattern in the middle of their CMAs. It is an interesting phenomena that all of the six CMAs have their hotspot of the greatest optimism located by the CMA boundary, indicated by the sentiment score estimated from the ratio of the sum of sentiment score over the sample size, which are Vancouver (sentiment score=0.817), Edmonton (sentiment score=0.794), Calgary (sentiment score=0.252), Toronto (sentiment score=0.181), Montréal (sentiment score=0.133), and Ottawa - Gatineau (sentiment score=0.564). And the magnitude of the positiveness at the hot spot are also varying a lot among these CMAs. One explanation for this inner-city disparity is that people are enjoying open green space much more than the urban life during the pandemic. Almost all of the hot spots with the highest estimated sentiment scores are located in the provincial parks or natural attractions. Vancouver's only hot spot at the north-east corner is in the Pinecone Burke Provincial Park; Edmonton's most optimistic spot sits at the beach to the west of Webamun Lake; Toronto's most optimistic spot is surrounded by the Glen Major Forest; Montréal's most optimistic spot is next to the Gault Nature Reserve of McGill University (Réserve naturelle Gault de l'Université McGill); and the most optimistic spot of Ottawa-Gatineau locates in the Wildlife Reserve of Papineau-Labelle (Réserve faunique de Papineau-Labelle).

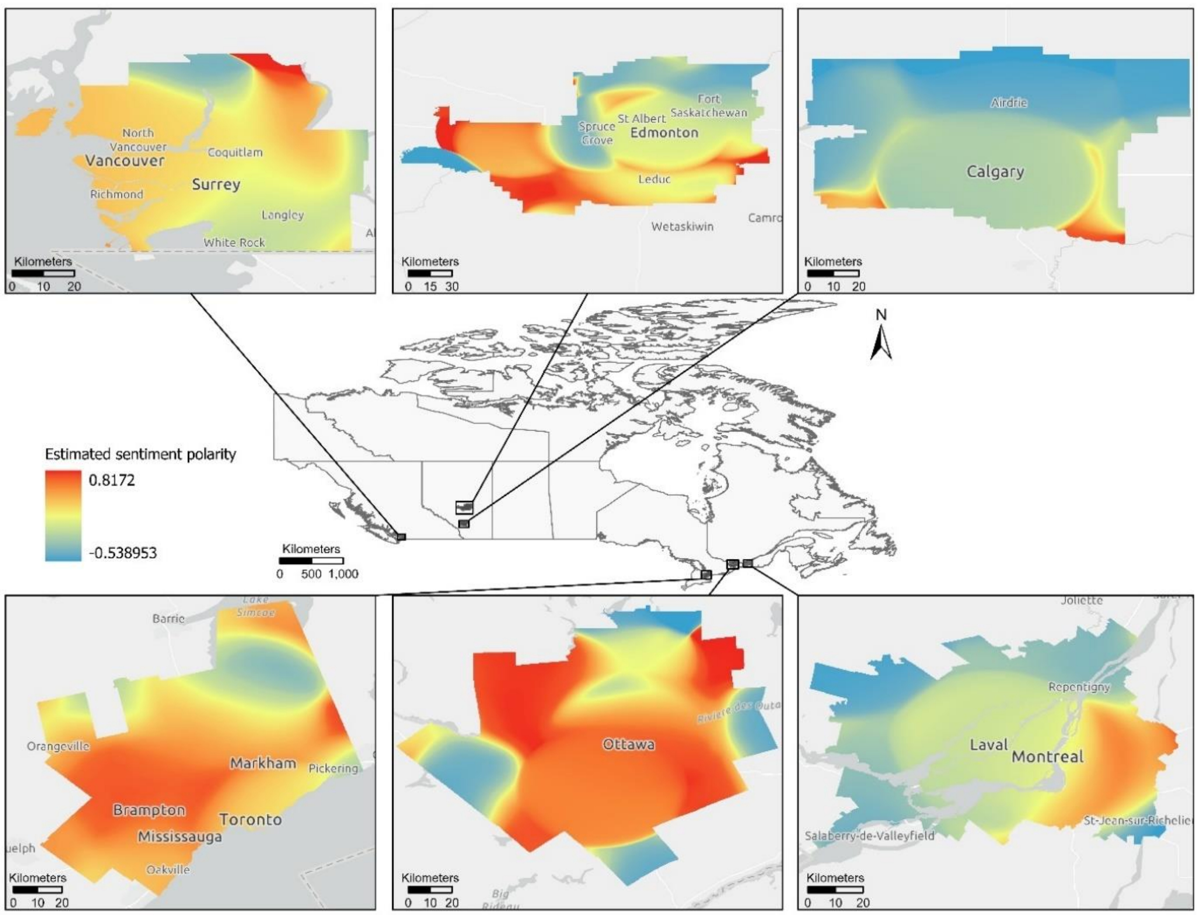


Figure 4.6 Kernel density ratio heatmap of positive and negative feelings across six major CMAs

To the opposite end of the sentiment polarity spectrum, the distribution of cold spots also varies a lot among the six areas. Some of the cities, Vancouver and Toronto have only one cold spot of pessimism of a clear oval shape, while the other municipalities have a much more irregular distribution of pessimism. Calgary and Montréal show a very uncommon pattern of pessimism spreading across more than two thirds of the area. In Calgary, the cold snap of pessimism flow heads from the north to the south, while in Montréal, it outflanks from the west and south-east to the middle. Cold spots spreaded in the outskirts area are observed in the other two CMAs, Ottawa-Gatineau.

Figure 4.7 depicts the trends of sentiment score along the pandemic timeline for each of the six major CMAs. In general, the Montréal CMA and the Calgary CMA have a substantially more depressing outlook than the other five metropolitan areas, particularly when we compare Montréal to its two neighbouring cities Toronto and Ottawa-Gatineau in the east coast. Beneath this fact, this might be a disparity caused by the different prevalence of English languages for Montréal in Québec and for other CMAs. The statistical summary of mean VADER score and NRCLex emotion percentages aggregated by Stage for Vancouver, Edmonton, Calgary, Toronto, Ottawa - Gatineau, and Montréal is given in Tables 4.2 to 4.7, respectively.

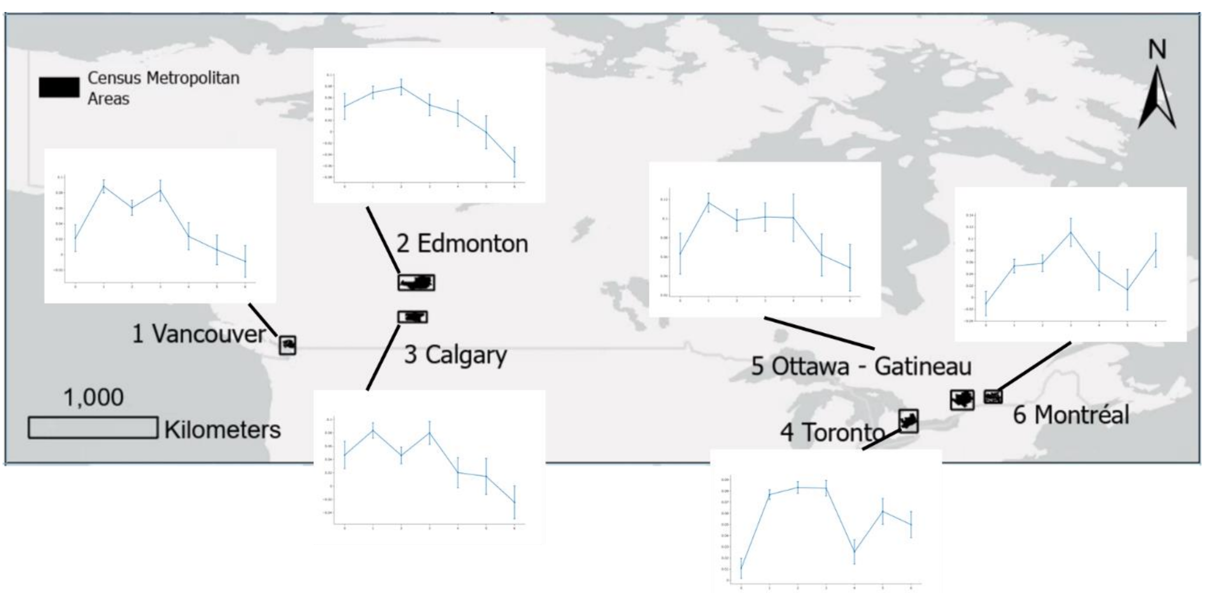


Figure 4.7 Mean sentiment score by stage in six major CMAs

Table 4.2 Mean VADER score and NRCLex emotion percentages by Stage in Vancouver

Wave	VADER score	Fear (%)	Anger (%)	Sadness (%)	Disgust (%)	Joy (%)	Anticipation (%)	Trust (%)	Surprise (%)
0	0.021 (0.009)	8.18 (0.27)	3.22 (0.15)	5.80 (0.21)	2.78 (0.15)	3.12 (0.16)	7.32 (0.34)	7.61 (0.31)	2.67 (0.18)
1	0.088 (0.004)	8.45 (0.13)	2.79 (0.07)	6.81 (0.11)	1.88 (0.06)	4.03 (0.08)	7.84 (0.16)	8.08 (0.15)	2.54 (0.08)
2	0.061 (0.005)	10.56 (0.15)	2.83 (0.07)	9.01 (0.13)	1.94 (0.06)	3.38 (0.08)	7.27 (0.17)	8.20 (0.17)	2.70 (0.10)
3	0.083 (0.007)	9.72 (0.20)	3.31 (0.12)	8.18 (0.17)	1.87 (0.08)	3.88 (0.12)	7.60 (0.23)	6.65 (0.19)	2.80 (0.11)
4	0.024 (0.009)	9.81 (0.25)	3.28 (0.15)	8.36 (0.23)	2.40 (0.12)	3.12 (0.14)	6.34 (0.27)	7.67 (0.28)	2.12 (0.13)
5	0.006 (0.010)	10.84 (0.28)	3.18 (0.16)	9.94 (0.26)	2.45 (0.13)	2.81 (0.14)	6.32 (0.29)	6.64 (0.28)	2.45 (0.17)
6	-0.009 (0.010)	11.55 (0.27)	3.65 (0.16)	10.16 (0.25)	2.63 (0.13)	2.95 (0.14)	6.10 (0.25)	6.65 (0.29)	2.54 (0.15)

Note. Bracket behind each value of the mean shows the standard error of the mean (SEM) that measures how far the sample mean is likely to be from the true population mean.

Table 4.3 Mean VADER score and NRCLex emotion percentages by Stage in Edmonton

Wave	VADER score	Fear (%)	Anger (%)	Sadness (%)	Disgust (%)	Joy (%)	Anticipation (%)	Trust (%)	Surprise (%)
0	0.044 (0.012)	7.97 (0.39)	2.77 (0.20)	5.45 (0.28)	2.52 (0.19)	3.38 (0.20)	8.76 (0.49)	8.39 (0.41)	2.92 (0.25)
1	0.069 (0.006)	8.30 (0.16)	2.98 (0.09)	7.07 (0.14)	1.92 (0.08)	3.72 (0.10)	7.69 (0.21)	8.10 (0.19)	2.44 (0.10)
2	0.079 (0.007)	9.49 (0.19)	3.04 (0.11)	8.57 (0.17)	2.25 (0.09)	3.45 (0.11)	6.59 (0.20)	7.88 (0.21)	3.02 (0.14)
3	0.047 (0.010)	8.98 (0.24)	3.28 (0.14)	8.68 (0.23)	2.24 (0.12)	3.40 (0.14)	7.17 (0.30)	6.23 (0.25)	2.51 (0.13)
4	0.032 (0.012)	10.45 (0.36)	3.08 (0.19)	8.64 (0.29)	2.42 (0.16)	2.97 (0.17)	5.26 (0.30)	8.28 (0.39)	2.00 (0.16)
5	-0.001 (0.015)	10.11 (0.41)	3.29 (0.23)	8.99 (0.37)	2.79 (0.22)	3.23 (0.24)	7.06 (0.46)	7.52 (0.47)	2.85 (0.29)
6	-0.053 (0.013)	11.70 (0.33)	3.69 (0.20)	10.42 (0.31)	2.94 (0.17)	3.21 (0.18)	5.91 (0.32)	7.49 (0.35)	2.26 (0.16)

Table 4.4 Mean VADER score and NRCLex emotion percentages by Stage in Calgary

Wave	VADER score	Fear (%)	Anger (%)	Sadness (%)	Disgust (%)	Joy (%)	Anticipation (%)	Trust (%)	Surprise (%)
0	0.047 (0.010)	8.00 (0.34)	2.71 (0.19)	5.72 (0.27)	2.92 (0.21)	3.39 (0.21)	6.97 (0.39)	8.83 (0.41)	3.08 (0.25)
1	0.084 (0.006)	8.80 (0.17)	3.23 (0.10)	7.47 (0.15)	2.09 (0.08)	3.97 (0.11)	7.63 (0.21)	7.99 (0.19)	2.56 (0.11)
2	0.046 (0.006)	10.80 (0.19)	3.12 (0.10)	8.85 (0.16)	2.26 (0.08)	3.46 (0.10)	7.47 (0.21)	7.51 (0.19)	2.45 (0.10)
3	0.080 (0.009)	9.65 (0.24)	3.04 (0.14)	8.45 (0.22)	2.10 (0.11)	3.32 (0.14)	7.04 (0.29)	6.61 (0.24)	2.35 (0.14)
4	0.020 (0.012)	9.54 (0.34)	3.71 (0.20)	8.34 (0.29)	2.59 (0.16)	3.04 (0.17)	6.01 (0.34)	7.94 (0.37)	2.98 (0.21)
5	0.015 (0.014)	9.54 (0.36)	3.74 (0.25)	8.31 (0.33)	2.60 (0.18)	3.62 (0.25)	5.57 (0.38)	6.91 (0.39)	2.50 (0.23)
6	-0.024 (0.013)	10.50 (0.34)	3.43 (0.20)	8.86 (0.30)	2.63 (0.17)	2.90 (0.17)	5.60 (0.30)	7.11 (0.36)	3.20 (0.26)

Table 4.5 Mean VADER score and NRCLex emotion percentages by Stage in Toronto

Wave	VADER score	Fear (%)	Anger (%)	Sadness (%)	Disgust (%)	Joy (%)	Anticipation (%)	Trust (%)	Surprise (%)
0	0.011 (0.005)	8.05 (0.15)	2.95 (0.08)	5.76 (0.11)	2.51 (0.08)	3.13 (0.08)	7.59 (0.18)	7.94 (0.17)	2.60 (0.09)
1	0.077 (0.002)	8.76 (0.07)	2.97 (0.04)	6.87 (0.06)	1.81 (0.03)	3.73 (0.04)	7.78 (0.08)	8.04 (0.08)	2.50 (0.04)
2	0.083 (0.003)	9.80 (0.08)	2.75 (0.04)	8.51 (0.07)	1.84 (0.03)	3.45 (0.05)	7.45 (0.10)	7.58 (0.09)	2.49 (0.05)
3	0.082 (0.004)	8.57 (0.10)	3.26 (0.06)	7.74 (0.09)	1.98 (0.05)	3.32 (0.06)	6.54 (0.11)	6.62 (0.10)	2.54 (0.06)
4	0.026 (0.006)	9.55 (0.15)	3.24 (0.09)	8.00 (0.13)	2.15 (0.07)	3.18 (0.09)	6.10 (0.16)	8.65 (0.19)	2.13 (0.08)
5	0.062 (0.006)	10.16 (0.17)	3.19 (0.10)	9.05 (0.16)	2.08 (0.07)	3.13 (0.09)	6.52 (0.19)	7.90 (0.20)	2.44 (0.11)
6	-0.009 (0.010)	11.55 (0.27)	3.65 (0.16)	10.16 (0.25)	2.63 (0.13)	2.95 (0.14)	6.10 (0.25)	6.65 (0.29)	2.54 (0.15)

Table 4.6 Mean VADER score and NRCLex emotion percentages by Stage in Ottawa-Gatineau

Wave	VADER score	Fear (%)	Anger (%)	Sadness (%)	Disgust (%)	Joy (%)	Anticipation (%)	Trust (%)	Surprise (%)
0	0.063 (0.011)	7.94 (0.35)	2.58 (0.18)	5.69 (0.26)	2.44 (0.18)	3.12 (0.19)	8.19 (0.45)	9.11 (0.43)	2.36 (0.20)
1	0.117 (0.005)	8.79 (0.15)	2.96 (0.09)	7.01 (0.13)	1.74 (0.06)	3.98 (0.10)	7.35 (0.17)	8.85 (0.18)	2.42 (0.09)
2	0.098 (0.006)	10.09 (0.18)	2.68 (0.09)	8.87 (0.16)	1.71 (0.07)	3.55 (0.10)	7.67 (0.20)	7.54 (0.18)	2.53 (0.11)
3	0.102 (0.008)	9.73 (0.22)	2.76 (0.11)	7.89 (0.19)	1.51 (0.08)	3.49 (0.12)	7.23 (0.24)	7.26 (0.22)	2.53 (0.13)
4	0.101 (0.013)	9.87 (0.33)	3.09 (0.20)	8.92 (0.32)	1.77 (0.13)	3.70 (0.20)	6.55 (0.33)	7.89 (0.37)	2.09 (0.17)
5	0.062 (0.011)	10.55 (0.31)	3.28 (0.18)	8.41 (0.28)	1.95 (0.13)	2.74 (0.16)	6.67 (0.37)	7.24 (0.35)	2.81 (0.26)
6	0.049 (0.012)	12.00 (0.34)	3.18 (0.18)	10.61 (0.31)	2.31 (0.16)	2.88 (0.17)	6.20 (0.34)	7.33 (0.37)	2.16 (0.16)

Table 4.7 Mean VADER score and NRCLex emotion percentages by Stage in Montréal

Wave	VADER score	Fear (%)	Anger (%)	Sadness (%)	Disgust (%)	Joy (%)	Anticipation (%)	Trust (%)	Surprise (%)
0	-0.010 (0.011)	7.79 (0.36)	3.42 (0.22)	5.24 (0.26)	3.38 (0.21)	2.80 (0.19)	7.47 (0.45)	7.25 (0.44)	2.89 (0.24)
1	0.054 (0.006)	8.28 (0.19)	3.17 (0.12)	6.73 (0.15)	1.92 (0.08)	3.55 (0.11)	6.72 (0.21)	7.61 (0.22)	2.32 (0.10)
2	0.058 (0.007)	9.01 (0.23)	3.19 (0.15)	7.49 (0.19)	2.04 (0.11)	3.44 (0.13)	6.47 (0.26)	6.37 (0.24)	2.23 (0.13)
3	0.111 (0.012)	9.32 (0.35)	2.59 (0.19)	8.54 (0.33)	1.61 (0.15)	3.75 (0.22)	7.64 (0.45)	7.47 (0.39)	2.79 (0.25)
4	0.045 (0.016)	9.40 (0.44)	3.45 (0.30)	7.99 (0.40)	2.51 (0.23)	3.18 (0.26)	5.38 (0.46)	7.82 (0.50)	2.53 (0.38)
5	0.013 (0.018)	10.81 (0.51)	2.98 (0.27)	10.16 (0.51)	1.96 (0.22)	2.75 (0.27)	6.34 (0.52)	6.65 (0.61)	2.76 (0.33)
6	0.080 (0.015)	11.50 (0.47)	2.89 (0.23)	9.91 (0.40)	1.60 (0.17)	2.64 (0.21)	7.91 (0.59)	6.17 (0.43)	2.41 (0.24)

4.5 General Discussion

This study analyzes GeoTweets posted between January 1, 2020 and December 31,2022 in order to extract and evaluate the mental health dynamics during COVID-19 in Canada. The specific methods include tracking spatiotemporal patterns of public sentiment and emotion throughout the Canadian pandemic timeline, modelling the trending Twitter keywords, as well as examining the potential emotional triggers for the general public during the pandemic. As a summary, the stage mean sentimental score surged rapidly from 0.008, the feet of the mountain, in the Stage 0 to the peak of 0.091 in Stage 3 and then declined through almost every stage to 0.015, back to the feet of the mountain, in the Stage 6, revealing an optimistic first and then pessimistic trend along the timeline. Among the seven stages of the pandemic in Canada, only two stages show a statistically significant increase in sentiment polarity than the last stage; the sentiment score of the most recent stage decreased by 66.72% of the previous stage, sending a strong warning signal in mental health.

Likewise, the results of emotion classification align with this worsening tendency. Out of the eight different emotions identified by NRCLex, fear and sadness consistently rank first and second, respectively, and continue to increase through the study's ending, in the post-pandemic stage. In contrast, trust and anticipation, the emotion taking turn to occupy the third and fourth position, has been shrinking since November 2021, the end of the fourth wave stage, all the way towards the future.

Next, if you switch the viewpoint from national level to the city level at the six Canadian CMAs with most population, you can see that five out of the six CMAs follow a coarse rise-fall-rise-fall-fall-fall pattern of sentiment polarity from Stage 0 to Stage 6 in Figure 4.7. Among them, three show a bimodal trend that peaks at Stage 1 and 3, while Edmonton reaches the highest optimism right at the middle of Stage 1 and 3, displaying a lightly different unimodal shape. Toronto has been one stage delayed to the majority, peaking at Stage 3 and 5 instead. In contrast, Montréal moves mostly in a counter direction characterized by rise-flat-rise-fall-fall-rise. With that said, the five CMAs including Vancouver, Calgary, Edmonton, Toronto and Ottawa-Gatineau are signaling a downward inclination of sentiment positivity towards the future and all six CMAs are expressing an exhilarating mental health status in terms of the expression of fear, sadness, anger and disgust over trust, anticipation, joy and surprise in pandemic discussion on Twitter.

Moreover, there is a revelatory parallel between the findings in this thesis and the estimated national trends in COVID-19 vaccine hesitancy in Canada revealed by Lavoie et al. (2022) as we perceive the sentiment dynamics from the angle of public vaccination acceptability. In this design, Lavoie et al. (2022) recruited a total of 15,019 Canadian adult participants in five successive cross-sectional representative surveys between April 2020 and March 2021, which overlaps the time period from Stage 1 to Stage 3 in the paradigm of this thesis. According to their analysis, percent of participants admitting vaccine hesitancy peaked in the middle of the pandemic at 52.9% in survey 3 (November 2020, two months before the peak of the second wave, Stage 2) and were lowest at 36.8% in survey 1 (April 2020, the start of the first wave, Stage 1) and 36.9% in survey 5 (March 2021, just after the start of the third wave, Stage 3). This N shape trend on vaccine hesitancy showing that people are the most reluctant to get vaccination despite the availability of vaccination services during Stage 2 matches the U shape trend on our estimated sentiment polarity indicating that people were most pessimistic towards the pandemic during Stage 2.

From another perspective, if you move the focus to the Freedom Convoy event, it is noticeable that a local minimum of monthly sentiment score from the mid-fall 2021 to the late spring 2022 appeared in February 2022, which overlapped the ending of Freedom Convoy (January 15 to February 14, 2022). This could be an implication of a negative reaction from the public to the new vaccine mandate released on January 15, 2022 which demand that all essential service providers, including truck drivers, must be vaccinated to enter the country and was the direct trigger of the Freedom Convoy, and the government response during the Freedom Convoy. Based on the investigations of Huang et al. (2022a), public's negative attitudes were most concentrated over the topic "vaccine mandates" out of the five most discussed topics about the Freedom Convoy, with other four topics being "convoy support", "fundraising", "police activities" and "Trudeau".

4.6 Chapter Summary

This chapter reports the outcome of sentiment analysis, emotion classification, keyword analysis and kernel density heatmap. We also explore the possible association between the observed patterns in sentiment polarity and crucial pandemic events such as government recognized the COVID-19 pandemic as emergency, vaccines became available in Canada, public's vaccine hesitancy changed, as well as government imposed stricter pandemic policies. As an overview, the stage mean sentimental score surged rapidly from 0.008, the feet of the mountain, in the pre-pandemic stage to the peak of 0.091 in the stage of third wave and then declined through almost every stage to 0.015, back to the feet of the mountain, in the current post-pandemic stage, revealing an optimistic first and then pessimistic trend in mental health along the timeline. In addition, fear and sadness consistently rank first and second, respectively, out of the eight different emotions identified by NRCLex and continue to increase through the post-pandemic stage. Simultaneously, the five CMAs including Vancouver, Calgary, Edmonton, Toronto and Ottawa-Gatineau are signaling a downward inclination of sentiment positivity towards the future and all six CMAs are showing an exhilarating mental health status in terms of the expression of fear, sadness, anger and disgust over trust, anticipation, joy and surprise in pandemic discussion on Twitter.

Chapter 5
Conclusions and Recommendations

5.1 Conclusions

This thesis explored and analyzed the evolution of pandemic as a stressor on mental health in Canada from January 2020 to December 2022. Machine learning algorithms including sentiment analysis models and emotion classification models as well as geospatial mapping method including kernel density mapping were deployed. Through monitoring the sentiment polarity dynamics, emotion trends, and changes in keywords being discussed on Twitter during the full three-year span, we were able to address some key questions about the spatiotemporal development of mental health signals.

First, to answer the research question of what the social media data informs us about the shifting in mental health condition along the Canadian pandemic timeline, our results reveal that the overall sentiment and emotion composition was the most optimistic during the early phase of the pandemic, from the early spring of 2020 to the summer of 2021, and turned to decline from then to the end of 2022, sending a warning signal in public mental health.

For the second research question of what critical events are driving the trends in mental health, we identify several driving events, which are the declaration of state of emergency in March 2020, the peak of vaccine hesitancy in November 2020, the release of new vaccine mandate in January 2022 and the Freedom Convey lasting from January to February 2022.

The findings of this study also indicate that there is an observable geospatial disparity in the shifting patterns and the overall mental health levels between Montréal, a French-dominant region, and Vancouver, Calgary, Edmonton, Toronto, and Ottawa-Gatineau, which are English-dominant or bilingual regions. This is in response to the third research question about the geospatial heterogeneity of mental health in Canada. Also, along with a delayed period of peaks and bottoms in sentiment polarity, Edmonton and Toronto are displaying a little different mood than the other English-speaking cities.

Last but not the least, responding to the fourth research question, this thesis proposes two government actions, promoting education on the importance of vaccine behaviours and rebalancing of the COVID-19 restrictions, as two critical public health strategies for boosting public confidence regarding the pandemic and rebuilding psychological resilience in the current post-pandemic era.

5.2 Public Health Implications

It is also worth noticing that in today's post-pandemic stage, as the global economy starts to recover and the number of cases becomes gradually under control with the availability of vaccines, the public psychological condition is not lifting as fast as the economy and the physical health. This study was the first work tracking the long-term mental health of Canada as a country during the pandemic and further supported the conclusion that the sentiment trend is not walking towards the best possible future, as previously indicated by studies in other areas of the world (Guntuku et al., 2020; Leng et al., 2021; Valdez et al., 2020; Wang et al., 2022; Yin et al., 2020).

One important implication is enhancing the quality and effectivity of vaccine education to the public. As revealed by this study and Lavoie et al. (2022), a worsening trend in mental health is echoing with a surging degree of vaccine hesitancy in the timeline. Therefore, spreading awareness of the value of preventive behaviours in minimizing the spread of viruses and avoiding significant harm from infection will not only assist solve the problem of vaccine hesitancy but also improve the overall situation of mental health.

Another implication is rebalancing and lifting the pandemic restrictions. Both the national-wide and most of the region-specific results exhibit a downward tendency in the mental health condition of post-pandemic stage, demanding the relevant sectors to seriously weigh the benefits against drawbacks of the current COVID-19 policies and reconsider lifting the controls. On the one hand, the benefits of strict pandemic regulations have been diminishing. As advocated by experts in the medical field, Daria & Islam (2022), the present omicron wave of SARS-CoV-2 no longer has an impact on human health as strong as the early waves of other variants. Even though the COVID-19 pandemic may continue to be a recurrent disease worldwide, it is time to think about ending the age of extreme safety precautions to avoid and limit virus infections (Daria & Islam, 2022). On the other hand, the unintended consequences of pandemic restrictions that remained in place have emerged. Not only has the 2022 vaccine mandate been identified in this study and by Huang et al. (2022) as a potential stressor on Canadians' psychological wellbeing, but "Pandemic fatigue" has also been observed as a result of the never-ending regulations that intensify the pandemic's cloud over citizens and threaten social connections even when the varieties become less dangerous (Haktanir et al., 2022; Lowe et al., 2023). In conclusion, rebalancing the COVID-19 regulations at all levels of governments would be vital in the fight against pandemic aftereffects in Canada.

5.3 Limitations and Recommendations for Future Research

While the present studies are offering a new perspective, there are still a lot of unanswered questions regarding the evolution of mental health during the COVID-19 pandemic in Canada, which is currently an area of little research. The first question is, would the gap between sentimental responses from Montréal and other cities including Toronto and Vancouver narrow or even close if we gather tweets in both French and English? The answer to this question remains unclear.

The second question is, how can we increase our data representativeness and data coverage of different demographic and socioeconomic backgrounds when we are sampling from GeoTweets? This problem is particularly important when the underrepresented group of people might heavily overlap with the most affected group of population under the pandemic, minority population and children.

As pointed out in the statistical summary of COVID-19 infection in the Canadian household population (Bushnik et al., 2022), people in a visible minority group were almost twice as likely as those not in a visible minority group to have been infected with the COVID-19 pandemic. When we are conducting the data mining in Tweets in English, we are not able to observe the minority population in Canada who express their opinions in social media in Mandarin, Cantonese, Punjabi, Spanish, Arabic, Tagalog, Hindustani, Portuguese and others language of minority more often than in English. Besides, there are currently no acknowledged methods or standards for determining the race or ethnicity of Twitter users in the academic field, since even manual validation method could be subjective and flawed (Golder et al., 2022), leaving us a long way in approaching a balanced and representative ethnicity structure in social media-based mental health research.

Likewise, bias in the age hierarchy is also a critical aspect to account for when we sample our data from GeoTweets. As suggested by a recent focus study (Malik et al., 2021) and a county-level case study in the US (Jiang et al., 2019), the age group under 18 years old is actually the most unrepresented group in age structure in Twitter due to their relative lack of access to social media compared to adults. Minors were one of the most vulnerable populations, as demonstrated in the statistical overview of COVID-19 infection in the Canadian household population (Bushnik et al., 2022). They were nearly twice as likely as people of older ages to have contracted COVID-19, and their voices should be considered in the study of maintaining and monitoring public psychological health. Future works that engage a wider population with a more diversified range of languages spoken, demographic structure and social media usage habits would take a larger step towards tackling these remaining challenges in the country-level mental health monitoring in Canada.

In addition to improving the ability to generalize the outcome to the general public, we would like to provide a couple of ideas to optimize the performance of sentiment analysis on tweets. Employing feature ensemble models to embed the original tweets and then feeding embedded tweets to sentiment analysis algorithms can help enhance the accuracy of sentiment classification in terms of F1 score, particularly for tweets with fuzzier emotions (Phan et al., 2020). Another way for researchers to strengthen such analysis is to implement more than one sentiment categorization models to estimate a common population parameter (i.e., happiness level, depression level, fear level, etc.) and study the correlation between those results.

References

Ajantha Devi, V., & Nayyar, A. (2021). Evaluation of geotagging twitter data using sentiment analysis during COVID-19. *Lecture Notes in Networks and Systems*, 601–608. https://doi.org/10.1007/978-981-15-9689-6_65

Atillo, G. N. A. (2021). Sentiment-Topic Dynamics on COVID-19 post-traumatic stress disorder using online published researches. *Asian Journal of Research in Education and Social Sciences, 3*(4), 28-35. https://doi.org/10.55057/ajress.2021.3.4.3

Banholzer, N., Van Weenen, E., Lison, A., Cenedese, A., Seeliger, A., Kratzwald, B., ... & Vach, W. (2021). Estimating the effects of non-pharmaceutical interventions on the number of new infections with COVID-19 during the first epidemic wave. *PLoS One*, 16(6), e0252827.

Birjali, M., Kasri, M., & Beni-Hssane, A. (2021). A comprehensive survey on sentiment analysis: Approaches, challenges and trends. *Knowledge-Based Systems*, 226, 107134.

Bolancé, C., Guillén, M., & Nielsen, J. P. (2008). Inverse beta transformation in kernel density estimation. *Statistics & Probability Letters, 78*(13), 1757–1764. https://doi.org/10.1016/j.spl.2008.01.028

Bushnik, T., Earl, S., Clark, J., & Cabot, J. (2022). COVID-19 infection in the Canadian household population. *Health Reports, 33*(4), 24-33.

Camus, A. (1991). *The Plague*. (S. Gilbert, Trans.). Vintage. https://doi.org/10.1604/9780679720218

Canada: metropolitan area population 2022 | Statista. (2023, January). Retrieved January 30, 2023, from https://www.statista.com/statistics/443749/canada-population-by-metropolitan-area/

Census of Population. (2023, March). Retrieved January 30, 2023, from https://www12.statcan.gc.ca/census-recensement/index-eng.cfm

Chen, Y. C. (2017). A tutorial on kernel density estimation and recent advances. *Biostatistics & Epidemiology, 1*(1), 161–187. https://doi.org/10.1080/24709360.2017.1396742

Daria, S., & Islam, M. R. (2022). The SARS-CoV-2 omicron wave is indicating the end of the pandemic phase but the COVID-19 will continue. *Journal of Medical Virology*, 94(6), 2343.

Dixon. (2022, December). Twitter: number of users worldwide 2019-2024 | Statista. Retrieved January 30, 2023, from https://www.statista.com/statistics/303681/twitter-users-worldwide/

Duong, T., & Hazelton, M. L. (2005). Cross-validation bandwidth matrices for multivariate kernel density estimation. *Scandinavian Journal of Statistics*, *32*(3), 485–506. https://doi.org/10.1111/j.1467-9469.2005.00445.x

Drias, Y., & Drias, H. (2021). Sentiment evolution analysis and association rule mining for COVID-19 tweets. *Journal of Digital Art & Humanities*, *2*(2), 3–21. https://doi.org/10.33847/2712-8148.2.2_1

Frank, M. R., Mitchell, L., Dodds, P. S., & Danforth, C. M. (2013). Happiness and the patterns of life: A study of geolocated tweets. *Scientific Reports*, *3*(1), 2625. https://doi.org/10.1038/srep02625

Galesic, M., Bruine de Bruin, W., Dalege, J., Feld, S. L., Kreuter, F., Olsson, H., . . . van der Does, T. (2021). Human social sensing is an untapped resource for computational social science. *Nature*, *595*(7866), 214–222. https://doi.org/10.1038/s41586-021-03649-2

Golder, S., Stevens, R., O'Connor, K., James, R., & Gonzalez-Hernandez, G. (2022). Methods to establish race or ethnicity of twitter users: Scoping review. *Journal of Medical Internet Research*, *24*(4), e35788.

Guntuku, S. C., Sherman, G., Stokes, D. C., Agarwal, A. K., Seltzer, E., Merchant, R. M., & Ungar, L. H. (2020). Tracking mental health and symptom mentions on twitter during COVID-19. *Journal of General Internal Medicine*, *35*(9), 2798–2800. https://doi.org/10.1007/s11606-020-05988-8

Guo, W., Gupta, N., Pogrebna, G., & Jarvis, S. (2016). Understanding happiness in cities using twitter: Jobs, children, and transport. In: *Proceedings of 2016 IEEE International Smart Cities Conference (ISC2)*. DOI: 10.1109/ISC2.2016.7580790.

Malik, M., Lamba, H., Nakos, C., & Pfeffer, J. (2015). Population bias in geotagged tweets. In: *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 9, No. 4, pp. 18-27.

Haktanir, A., Can, N., Seki, T., Kurnaz, M. F., & Dilmaç, B. (2022). Do we experience pandemic fatigue? current state, predictors, and prevention. *Current Psychology*, 41(10), 7314-7325.

Hu, T., Wang, S., Luo, W., Zhang, M., Huang, X., Yan, Y., . . . Li, Z. (2021). Revealing public opinion towards COVID-19 vaccines with twitter data in the United States: Spatiotemporal perspective. *Journal of Medical Internet Research*, *23*(9). https://doi.org/10.2196/30854

Huang, S. H., Tsao, S. F., Chen, H., Bin Noon, G., Li, L., Yang, Y., & Butt, Z. A. (2022a). Topic modelling and sentiment analysis of tweets related to freedom convoy 2022 in Canada. *International Journal of Public Health*, *67*. https://doi.org/10.3389/ijph.2022.1605241

Huang, X., Wang, S., Zhang, M., Hu, T., Hohl, A., She, B., . . . Li, Z. (2022b). Social media mining under the COVID-19 context: Progress, challenges, and opportunities. *International Journal of Applied Earth Observation and Geoinformation*, *113*, 102967. https://doi.org/10.1016/j.jag.2022.102967

Hollander, J. B., & Hartt, M. D. (2018). Big data and shrinking cities: How Twitter can help determine urban sentiments. In: *Big Data for Regional Science*, pp. 265-273.

Hutto, C., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. In: *Proceedings of the International AAAI Conference on Web and Social Media*, *8*(1), 216–225. https://doi.org/10.1609/icwsm.v8i1.14550

Jaidka, K., Giorgi, S., Schwartz, H. A., Kern, M. L., Ungar, L. H., & Eichstaedt, J. C. (2020). Estimating geographic subjective well-being from Twitter: A comparison of dictionary and data-driven language methods. In: *Proceedings of the National Academy of Sciences*, *117*(19), 10165–10171. https://doi.org/10.1073/pnas.1906364117

Jiang, Y., Li, Z., & Ye, X. (2019). Understanding demographic and socioeconomic biases of geotagged Twitter users at the county level. *Cartography and Geographic Information Science, 46*(3), 228-242.

Karami, A., Kadari, R. R., Panati, L., Nooli, S. P., Bheemreddy, H., & Bozorgi, P. (2021). Analysis of geotagging behavior: Do geotagged users represent the twitter population? *ISPRS International Journal of Geo-Information, 10*(6), 373. https://doi.org/10.3390/ijgi10060373

Lavoie, K., Gosselin-Boucher, V., Stojanovic, J., Gupta, S., Gagné, M., Joyal-Desmarais, K., ... & Bacon, S. (2022). Understanding national trends in COVID-19 vaccine hesitancy in Canada: results from five sequential cross-sectional representative surveys spanning April 2020–March 2021. *BMJ Open*, *12*(4), e059411.

Leng, Y., Zhai, Y., Sun, S., Wu, Y., Selzer, J., Strover, S., & Ding, Y. (2021). Misinformation during the COVID-19 outbreak in China: cultural, social and political entanglements. *IEEE Transactions on Big Data*, *7*(1), 69–80. https://ieeexplore.ieee.org/abstract/document/9340553

Li, M., Ch'ng, E., Chong, A., & See, S. (2017). The new eye of smart city: novel citizen sentiment analysis in twitter. In: *Proceedings of 2016 International Conference on Audio, Language and Image Processing (ICALIP)*. pp. 557-562.

Li, Z., & Wang, X. (2021). Research on Sentiment Analysis Model Due to COVID-19 Using Social Media. *Journal of Physics: Conference Series*, *1856*(1), 012064. https://doi.org/10.1088/1742-6596/1856/1/012064

Liu, Y., Liu, X., Gao, S., Gong, L., Kang, C., Zhi, Y., . . . Shi, L. (2015). Social sensing: A new approach to understanding our socioeconomic environments. *Annals of the Association of American Geographers*, *105*(3), 512–530. https://doi.org/10.1080/00045608.2015.1018773

Lowe, C., Rafiq, M., MacKay, L. J., Letourneau, N., Ng, C. F., Keown-Gerrard, J., ... & Ross, K. M. (2023). Impact of the COVID-19 pandemic on Canadian social connections: A thematic analysis. *Journal of Social and Personal Relationships*, 40(1), 76-101. https://doi.org/10.1177/02654075221113365

Luscombe, A., & McClelland, A. (2020). Policing the pandemic: Tracking the policing of COVID-19 across Canada. https://www.academia.edu/download/78531445/Policing_the_Pandemic_White_Paper_April_9_2020.pdf

Mitchell, L., Frank, M. R., Harris, K. D., Dodds, P. S., & Danforth, C. M. (2013). The geography of happiness: Connecting twitter sentiment and expression, demographics, and objective characteristics of place. *PloS One*, *8*(5), e64417.

Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing A Word-Emotion Association Lexicon. *Computational Intelligence*, *29*(3), 436–465. https://doi.org/10.1111/j.1467-8640.2012.00460.x

Niu, H., & Silva, E. A. (2020). Crowdsourced data mining for urban activity: Review of data sources, applications, and methods. *Journal of Urban Planning and Development*, *146*(2). https://doi.org/10.1061/(asce)up.1943-5444.0000566

Oesper, L., Merico, D., Isserlin, R., & Bader, G. D. (2011). WordCloud: A Cytoscape plugin to create a visual semantic summary of networks. *Source Code for Biology and Medicine*, *6*(1). https://doi.org/10.1186/1751-0473-6-7

Pandian, A. P. (2021). Performance evaluation and comparison using deep learning techniques in sentiment analysis. *Journal of Soft Computing Paradigm*, *3*(02), 123-134.

Phan, H. T., Tran, V. C., Nguyen, N. T., & Hwang, D. (2020). Improving the performance of sentiment analysis of tweets containing fuzzy sentiment using the feature ensemble model. *IEEE Access, 8*, 14630-14641.

Powell, L. M., Rebman, C.M., Dempsey, A., Myers, C.J. (2021). Using sentiment analysis to measure emotional toxicity of social media data during the COVID pandemic. *Issues in Information Systems*, *22*(1), 200-214. https://doi.org/10.48009/1_iis_2021_200-214

Quercia, D., Ellis, J., Capra, L., & Crowcroft, J. (2012). Tracking" gross community happiness" from tweets. In: *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, 965-968. https://doi.org/10.1145/2145204.2145347

Quercia, D., O Seaghdha, D., & Crowcroft, J. (2021). Talk of the city: Our tweets, our community happiness. In: *Proceedings of the International AAAI Conference on Web and Social Media*, *6*(1), 555–558. https://doi.org/10.1609/icwsm.v6i1.14302

Rosenblatt, M. (1969). Conditional probability density and regression estimators. *Multivariate Analysis II*, Academic Press, New York, pp. 25-31.

Sampogna, G., Pompili, M., & Fiorillo, A. (2021). Mental health in the time of COVID-19 pandemic: A worldwide perspective. *International Journal of Environmental Research and Public Health*, *19*(1), 161. https://doi.org/10.3390/ijerph19010161

Search Tweets version comparison. (n.d.). Retrieved January 30, 2023, from https://developer.twitter.com/en/docs/twitter-api/tweets/search/migrate

Tsao, S. F., Chen, H., Tisseverasinghe, T., Yang, Y., Li, L., & Butt, Z. A. (2021). What social media told us in the time of COVID-19: a scoping review. *The Lancet Digital Health*, *3*(3), e175–e194. https://doi.org/10.1016/s2589-7500(20)30315-0

Twitter. (n.d.). Twitter API for Academic Research | Products. Retrieved January 30, 2023, from https://developer.twitter.com/en/products/twitter-api/academic-research

Twitter. (n.d.). Advanced filtering for geo data. Retrieved January 31, 2023, from https://developer.twitter.com/en/docs/tutorials/advanced-filtering-for-geo-data

Umair, A., Masciari, E., & Habib Ullah, M. H. (2021). Sentimental analysis applications and approaches during COVID-19: A survey. In: *Proceedings of 25th International Database Engineering & Applications Symposium*. https://doi.org/10.1145/3472163.3472274

Valdez, D., ten Thij, M., Bathina, K., Rutter, L. A., & Bollen, J. (2020). Social media insights into US mental health during the COVID-19 pandemic: Longitudinal analysis of twitter data. *Journal of Medical Internet Research*, *22*(12), e21418. https://doi.org/10.2196/21418

Wang, D., Szymanski, B. K., Abdelzaher, T., Ji, H., & Kaplan, L. (2019). The age of social sensing. *Computer*, *52*(1), 36–45. https://doi.org/10.1109/mc.2018.2890173

Wang, S., Huang, X., Hu, T., Zhang, M., Li, Z., Ning, H., . . . Li, X. (2022). The times, they are a-changin': tracking shifts in mental health signals from early phase to later phase of the COVID-19 pandemic in Australia. *BMJ Global Health*, *7*(1), e007081. https://doi.org/10.1136/bmjgh-2021-007081

Wang, S., Zhang, M., Huang, X., Hu, T., Li, Z., Sun, Q. C., & Liu, Y. (2022). Urban-regional disparities in mental health signals in Australia during the COVID-19 pandemic: a study via Twitter data and machine learning models. *Cambridge Journal of Regions*, Economy and Society, 15(3), 663-682.

Wang, S., Huang, X., Hu, T., She, B., Zhang, M., Wang, R., ... & Bao, S. (2023). A global portrait of expressed mental health signals towards COVID-19 in social media space. *International Journal of Applied Earth Observation and Geoinformation*, 116, 103160.

Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, 55(7), 5731-5780.

Węglarczyk, S. (2018). Kernel density estimation and its application. In: *Proceedings of ITM Web of Conferences, 23*, 00037. https://doi.org/10.1051/itmconf/20182300037

Weng, Z., & Lin, A. (2022). Public opinion manipulation on social media: Social network analysis of twitter bots during the COVID-19 pandemic. *International Journal of Environmental Research and Public Health*, *19*(24), 16376. https://doi.org/10.3390/ijerph192416376

World Health Organisation. (2020). Timeline: WHO's COVID-19 response. Retrieved January 10, 2023, from https://www.who.int/emergencies/diseases/novel-coronavirus-2019/interactive-timeline#event-72

Yadav, A., & Vishwakarma, D. K. (2020). Sentiment analysis using deep learning architectures: a review. *Artificial Intelligence Review*, *53*(6), 4335-4385.

Yang, X. (Ed.). (2021). Urban Remote Sensing. *Monitoring, Synthesis and Modeling in the Urban Environment*. Wiley-Blackwell.

Yin, H., Yang, S., & Li, J. (2020). Detecting topic and sentiment dynamics due to COVID-19 pandemic using social media. *Advanced Data Mining and Applications*, 610–623. https://doi.org/10.1007/978-3-030-65390-3_46

Appendix A
Supplemental Material

A.1 Examples of location information in GeoTweets

There are two types of geographical tags in GeoTweets. Users can embed their location in a tweet by either tagging a pair of precise coordinates or tagging a Twitter "place" object.

An example of a geotag as a pair of coordinates returned by Twitter API looks like:


```
  "coordinates": {
    "type": "Point",
    "coordinates": [
      -149.90629456,
      61.19710597
    ]
}
```


An example of a geotag as a place object returned by Twitter API looks like:


```
  "place": {
    "geo": {
      "type": " Feature ",
      "bbox": [-79.639319, 43.403221, -78.90582, 43.855401],
      "properties": {}
    },
    "id": " 3797791ff9c0e4c6 ",
    "place_type": " city ",
    "name": " Toronto",
    "full_name": " Toronto, Ontario",
    "country_code": "CA",
    "country": "Canada",
}
```