UNIVERSITY OF CYPRUS

DEPARTMENT OF COMPUTER SCIENCE

PRESENTATION PLAN

OF INDIVIDUAL DIPLOMA WORK

May 2024

Diploma Project

TRACKING EMOTIONAL SHIFTS: ANALYZING TWEET SENTIMENT PRE, DURING, AND POST COVID-19

Simona Ivanova Uteva

UNIVERSITY OF CYPRUS

DEPARTMENT OF COMPUTER SCIENCE

May 2024

UNIVERSITY OF CYPRUS

DEPARTMENT OF COMPUTER SCIENCE

Tracking Emotional Shifts: Analyzing Tweet Sentiment

Pre, During, and Post COVID-19

Simona Ivanova Uteva

Supervisor

George Pallis

The Individual Thesis was submitted in partial fulfillment of the requirements for obtaining the Computer Science degree of the Department of Computer Science of the University of Cyprus

May 2024

SUMMARY

This thesis investigates changes in public sentiment during the COVID-19 pandemic using Twitter data spanning from January 2018 to February 2023. The study focuses on global trends and conducts detailed case studies of the USA and Sweden to analyze how different national strategies influenced public emotions. The primary objectives were to analyze the evolution of public sentiment during the pandemic, identify trends, evaluate the impact of significant events such as lockdowns and vaccination rollouts on sentiment, perform a comparative analysis of sentiment across different countries, and conduct topic modeling to categorize tweets and analyze emotional fluctuations within these sectors.

Data was collected from Twitter and Our World in Data. The study employed sentiment analysis tools such as VADER, TextBlob and machine learning models like RoBERTa. The analysis included correlational studies, sentiment fluctuations over time, statistical testing, and topic modeling.

Key findings include that the pandemic significantly influenced emotions, with a notable increase in negative sentiments and a decrease in positive sentiments. In the USA, fear and sadness showed the biggest spikes at the pandemic's onset, with anger, offensive score, and hate metrics escalating during major global events. Increased hostility in social media interactions was noted, with significant rises in hate scores and offensiveness. In Sweden, distinct emotional patterns were influenced by the country's unique approach to managing the crisis. Unlike the USA, there was no substantial change in outward expressions of fear, anger, or offensive language. The study also conducted analysis based on topic modeling to further categorize and understand emotional fluctuations across different sectors.

This comprehensive analysis of public sentiment during the COVID-19 pandemic provides valuable insights into the dynamic and complex nature of public emotions in response to global crises.

ACKNOWLEDGEMENTS

First and foremost, I express my gratitude to my thesis supervisor, Professor George Pallis, for introducing me to this intriguing field and supporting me throughout my academic journey. I admire his commitment to providing the best learning experience and support for his students. I am also thankful to Dimosthenis Stephanis, whose assistance and contributions were essential in overcoming the challenges we faced. Their expertise and guidance made it possible to bring this project to fruition.

A special thank you goes to my classmate, Nicolas Theodosiou, who worked on this thesis alongside me. Collaborating with Nicolas, sharing ideas, and challenging each other's thinking have been instrumental to our success. I could not have asked for a better companion on this academic adventure.

I must also acknowledge my friends, who have been with me throughout this process. Their unquestionable support and belief in my abilities pushed me to always give my best. Thank you for being such try hards and making me one too.

Most importantly, no acknowledgment would be complete without mentioning his esteemed Fluffiness Mr. Manji. Despite his best efforts to always steer me away from completing this degree, he has failed. Nonetheless, he should be given a heartfelt thank you for never leaving my side throughout this whole journey.

Thank you all for the part you played in this.

CONTENTS

CHAPTER 1: INTRODUCTION

1.1 BACKGROUND

INTRODUCTION TO SENTIMENT ANALYSIS

Sentiment analysis is a field within natural language processing (NLP) that involves the study of sentiments and emotions expressed in text. It aims to evaluate the sentiment behind a piece of text, categorizing it as positive, negative, or neutral. This process often involves various techniques, ranging from simple rule-based approaches to more complex machine learning models. (IBM)

The large volume of data generated by users on social media platforms provides a rich source of information for sentiment analysis. Platforms like Twitter, Facebook, and Reddit have become popular tools for individuals to express their opinions and emotions, making them valuable resources for scientists.

Tools like VADER (Valence Aware Dictionary and Sentiment Reasoner), LDA (Latent Dirichlet Allocation), and TextBlob are usually used to analyze social media posts. These tools are designed to interpret the informal language used on social media and capture nuanced sentiments.

CONTEXT OF THE STUDY

The COVID-19 pandemic, which was declared such on March 11th 2020 by the World Health Organization, is perhaps the most important crisis of the 21st century. The pandemic has caused significant disruptions to daily life, economies, and healthcare systems worldwide. With the spread of the virus a lot of changes in daily life occurred, causing a variety of emotional shifts in the population. Governments implemented various measures to control the spread, such as lockdowns, social distancing, and vaccination campaigns, all of which significantly influenced public sentiment. (WHO, 2020)

Social media platforms, particularly Twitter, became key for public discourse during the pandemic. People used these platforms to share their experiences, opinions, and emotions. This massive influx of data presents an opportunity to apply sentiment analysis to understand how public sentiment evolved during different phases of the pandemic.

1.2 RESEARCH OBJECTIVES

PROBLEM STATEMENT

The COVID-19 pandemic has caused significant disruptions to daily life, economies, and healthcare systems worldwide, leading to profound shifts in public sentiment. This study addresses the need to understand the evolution of public sentiment during the COVID-19 pandemic. Specifically, by analyzing Twitter data, it seeks to explore how public emotions fluctuated in response to key events and government measures, worldwide as well as in specific countries such as the USA and Sweden. Understanding these emotional dynamics is essential for developing effective public health strategies and improving communication during global crises. This study aims to provide a thorough analysis of public sentiment during the COVID-19 pandemic, offering valuable insights into the complex nature of public emotions in response to major global challenges.

RESEARCH OBJECTIVES

- To analyze the evolution of public sentiment during the COVID-19 pandemic using social media data from Twitter.
    - Identify Sentiment Trends: Identify key trends and patterns in public sentiment over different phases of the pandemic.
    - Impact of Major Events: Evaluate the impact of significant events, such as lockdowns, vaccination rollouts, and other important events, on public sentiment.
    - Geographical Analysis: Perform a comparative analysis of sentiment across different countries and regions, focusing on the USA and Sweden as case studies.
    - Topic Modeling: To conduct topic modeling to categorize tweets into different sectors and analyze emotional fluctuations within these sectors.

By addressing these objectives, this study aims to provide a comprehensive understanding of public sentiment during the COVID-19 pandemic.

1.3 SIGNIFICANCE OF THE STUDY

This study contributes to the academic field by providing a comprehensive analysis of public sentiment during one of the most significant global crises of modern times. By utilizing advanced sentiment analysis techniques, including both traditional lexicon-based methods and machine learning models, this research advances the understanding of how public sentiment evolves in response to crises.

This document builds on existing literature by performing a comparative analysis of sentiment trends in the USA and Sweden offering valuable insights into how different health strategies can influence public sentiment. By using topic modeling to categorize tweets into different sectors and examining emotional fluctuations within these sectors, this study offers a detailed understanding of how different domains were perceived during the pandemic.

Furthermore, the study covers an extended timeframe from January 2018 to February 2023, allowing for the observation of long-term trends and the impact of major events over time.

1.4 SCOPE AND LIMITATIONS

This study aims to analyze public sentiment during the COVID-19 pandemic with the use of Twitter data spanning from January 2018 to February 2023. Twitter is used as the primary data source due to its popularity as a mode of discourse and large number of users worldwide. The study explores sentiment changes across the globe and looks in depth into the cases of the USA and Sweden due to their different responses to the pandemic. Methodologically, it combines traditional lexicon-based tools like VADER and TextBlob with advanced machine learning models like RoBERTa, to analyze sentiment. Topic modeling is also applied to categorize tweets by sectors such as health, politics, and finance, allowing for an exploration of emotional fluctuations within these areas. The time frame spans five years to cover sentiment trends before, during, and after the pandemic's peak phases.

This study offers significant contributions to both the academic and practical understanding of public sentiment during the COVID-19 pandemic. While it provides valuable insights, we should also acknowledge its limitations.

The reliance on Twitter data introduces potential biases, as the platform's users may not represent the general population. Additionally, the analysis is limited to English tweets and the accuracy of sentiment analysis tools like VADER, TextBlob, and RoBERTa. And while sentiment analysis provides valuable insights, it may lack the depth of understanding that qualitative analyses offer, potentially missing some contextual undertones. Lastly, despite covering a significant period, the changing nature of discourse on social media means that findings may not be applicable outside the analyzed timeframe. (Detailed analysis on the limitations can be found in chapter 3.3)

By recognizing these limitations, the study aims to present a nuanced analysis on the impact of COVID-19 in the next sections by diving deeper into pre-existing work, our methodology, and results in the following chapters.

CHAPTER 2: LITERATURE REVIEW

Sentiment analysis is key for understanding public reactions, particularly during crises like the COVID-19 pandemic. As the pandemic unfolded, sentiment analysis provided insights into the emotional responses of populations globally, reflecting shifts in public mood in real-time as new information and developments emerged.

2.1 EVOLUTION OF PUBLIC SENTIMENT DURING COVID-19

The COVID-19 pandemic has profoundly influenced public sentiment, with noticeable shifts that varied over the course of the health crisis. Various studies on this topic were conducted, each of which provided unique insights into these shifts. In the following section we explore some of the existing studies and their findings.

In an article titled "Public Perception of the COVID-19 Pandemic on Twitter: Sentiment Analysis and Topic Modeling Study", Sakun Boon-Itt and Yukolpat Skunkan provide a peek into the early stages of the COVID-19 pandemic. They gathered tweets from December 2019 to March 2020 and investigated public sentiment, which at the time reflected widespread fear, anxiety, and uncertainty.

They found that the initial reactions were heavily negative, as people were concerned with the virus's rapid spread and its implications on health and economic stability. However, later on in the pandemic, a shift toward positive sentiments was observed, particularly when discussing advancements in prevention and treatment. Despite these positive shifts, the predominant sentiment remained negative. This highlights the complexity of public emotions during such a global crisis, showing gradual but incomplete shift toward optimism as scientific and medical responses progressed. (Boon-Itt & Skunkan, 2020)

Similarly, a study by Chandrasekaran et al 2020, focusing on English language tweets which spanned from January 1 to May 9, 2020 found similar results. They found that at first, people were mainly worried about the virus spreading, its symptoms, and its socio-economic effects, driven by uncertainty and the rapid rise in cases. Over time however, discussions around prevention methods, government actions, and healthcare advancements began reflecting more positive tones. This transition in sentiment suggested a growing public confidence in the efficacy of measures and progress toward controlling the pandemic. (Chandrasekaran et al 2020)

Lockdown measures introduced another dimension to public sentiment during the pandemic. In her study 2023 study, Carolina Biliotti found that initially, these measures exacerbated feelings of uncertainty and negativity, particularly in health and political spheres. Restrictions increased public anxiety, adding to existing negative feelings in different areas. (Biliotti, 2023). This analysis demonstrated how public emotion can be significantly swayed by government policies and their immediate impact on daily life, showcasing the direct influence of lockdowns on public sentiment.

The development and dissemination of COVID-19 vaccines marked a pivotal shift in public sentiment, particularly on platforms like Twitter and several articles were published about it.

Sentiment analysis conducted by Joanne and Han Luy in 2021 revealed a notable increase in positive sentiments following significant vaccine-related announcements, such as the Pfizer vaccine's reported 90% efficacy. These announcements shifted the public's predominant emotion from fear to trust, reflecting growing confidence in scientific advancements and their potential for overcoming the pandemic (Luy & Garving, 2021).

A similar sentiment analysis of tweets from India concerning COVID-19 and vaccination efforts from Bharati Sanja Ainapure, used machine learning techniques to classify sentiments. The results indicated a predominance of positive sentiments towards vaccination efforts, particularly highlighted by discussions on vaccine efficacy and rollout, reflecting widespread trust and optimism. Conversely, negative sentiments were also noted, largely stemming from concerns over vaccine availability, as well as general anxiety regarding the ongoing pandemic. (Ainapure et al 2023)

Differences in sentiment were also evident towards different COVID-19 vaccines. For instance, Pfizer and Moderna vaccines experienced higher positive sentiments compared to others like Sinopharm and Covaxin. These variations were influenced by specific news events and global developments, reflecting the dynamic nature of public opinion in response to information about vaccine efficacy and safety (Mushtag et al 2022).

Another study from Adwita Arora in 2023 on sentiment analysis during the COVID-19 pandemic revealed significant fluctuations that corresponded with the virus's waves. Each wave of infections resulted in a downturn in public sentiment, reflecting growing concerns and anxieties. These sentiments varied regionally, with developed countries showing more negativity due to concerns over virus variants, while developing countries like India displayed more positive sentiments following successful vaccination efforts and declining case numbers (Arora et al, 2023,). These regional differences illustrate the impacts of local circumstances on public sentiment.

Prior to Arora’s work, a north American study from Zhang et al used tweets from February to October 2020 to perform a sentiment analysis from major north American cities before vaccine rollout. It indicated that public sentiment towards pandemic measures like masks, vaccines, and lockdowns shifted over time. Initially negative, public sentiment became more positive as people adjusted to new norms, though it dipped again with subsequent infection waves. Sentiment was generally mixed; masks and lockdown measures, showed negative sentiment indicative of public fatigue, while vaccine discussions showed more optimism or skepticism. Notably, sentiment differed between Canadian and U.S. cities, likely reflecting the different public health strategies and media influences. This study highlights how local conditions and policy changes can profoundly influence public sentiment. Finally, the study by Zhang found that there is a direct correlation between tweet engagement in the form of likes and retweets and its sentiment index. (Zhang et al 2020).

2.2 METHODOLOGICAL APPROACHES IN SENTIMENT ANALYSIS

The methodologies used in sentiment analysis of social media posts encompass a range of both traditional lexicon-based approaches and advanced machine learning techniques. These methods are crucial for extracting and interpreting sentiments from informal and often abbreviated language that is found in social media texts like tweets.

A lot of studies such as those from Mushtaq, Ainapure, Chandrasekaran, Zhang and Arora utilize VADER (Valence Aware Dictionary and sEntiment Reasoner) and TextBlob for their analyses. VADER is renowned for its effectiveness with social media language, utilizing a lexicon of sentiment-related words and accounting for the context in which terms appear to determine sentiment polarity. TextBlob on the other hand offers a straightforward API for handling common natural language processing tasks, aiding in the analysis of text polarity, which is vital for sentiment analysis. (Mushtag et al 2022, Ainapure et al 2023, Chandrasekaran et al 2020, Zhang et al 2020, Arora et al 2023)

In the work by Zhang et al, the RoBERTa model -an optimized BERT pretraining approach tailored for sentiment analysis – was also employed. This model captures contextual relationships between words using deep learning, therefore providing a sophisticated understanding of how public emotions fluctuate over time and in response to specific pandemic-related events. (Zhang et al 2020)

Another common model was Latent Dirichlet Allocation (LDA), employed in the studies of Lyu, Boon-Itt, and Chandrasekaran, which allows for identification of common themes or topics within the textual data. LDA is a form of topic modeling that groups similar words into topics, proving invaluable in categorizing and summarizing large datasets for sentiment analysis. This approach helps reveal the structure of sentiment data, revealing the primary topics of discussion and their associated sentiments. (Lyu et al 2021, Boon-Itt et al 2020, Chandrasekaran 2020)

Additional sophisticated tools were incorporated in other studies, such as NCRLex, which offer a comprehensive lexicon including a range of emotions for a more detailed analysis beyond the simple positive, negative, and neutral categories. This enhances the granularity of sentiment and emotional analysis, mapping complex emotional states expressed in social media posts. (Ainapure et al 2023, Zhang et al 2020)

The study of Ainapure et al delved into deep learning architectures like Bidirectional Long Short-Term Memory (Bi-LSTM) and Gated Recurrent Units (GRU). These techniques are best at processing sequential data, capturing information from both previous and future contexts, which is great for a comprehensive understanding of sentiments expressed over time or across conversations. (Ainapure et al 2023)

Lastly, all of the above the studies applied advanced statistical techniques to analyze their results, emphasizing the importance of robust data analysis methods to derive meaningful insights from large volumes of data.

2.3 CHALLENGES AND LIMITATIONS

In summarizing the main challenges and limitations faced by the studies on public sentiment analysis during the COVID-19 pandemic, several recurrent themes emerge across the different articles:

SOCIAL MEDIA BIAS

A common limitation cited across multiple studies is the bias inherent in social media data. Twitter users may not be representative of the general population, potentially skewing sentiment analysis results. This demographic bias can lead to overrepresentation of younger, more tech-savvy individuals, which might not accurately reflect the sentiments of older or less digitally educated demographics.

GEOGRAPHICAL AND LANGUAGE LIMITATIONS

Many studies note challenges related to geographical and language biases. Sentiments analyzed were often only from tweets in English or from users in specific geographic locations, which could overlook the sentiments of non-English speaking populations or those outside major urban centers. This limitation restricts the studies' ability to portray a comprehensive picture of global sentiments.

TEMPORAL LIMITATIONS

Several studies face challenges due to the specific time frames of data collection. Public sentiment can fluctuate significantly, especially in response to new information or developments regarding the pandemic. Studies covering only specific periods may not capture these changes, limiting their relevance over time.

ACCURACY OF SENTIMENT ANALYSIS TOOLS

The accuracy of sentiment analysis tools like VADER and TextBlob is questioned, particularly their ability to capture complex emotions, sarcasm, and the nuances of human language. These models, despite being state-of-the-art, are not infallible and can introduce errors, which can lead to misclassification of sentiments or misinterpretation of data, particularly when dealing with informal and diverse language as found in tweets.

CHALLENGES WITH DATA COLLECTION AND QUALITY

The quality of data collected from social media can be noisy and informal, which poses significant challenges in accurately interpreting and processing text. Additionally, the articles of Boon-Itt and Arora rely on specific keywords and hashtags for data collection that could result in incomplete datasets, potentially biasing the results.

ETHICAL AND PRIVACY CONSIDERATIONS

Several studies acknowledge ethical challenges, especially regarding the privacy of individuals and the potential misuse of sentiment analysis results. The ethical implications of aggregating and analyzing user-generated content are critical considerations that require careful handling.

These challenges highlight the complexities involved in using sentiment analysis for studying public opinions on social media. They underscore the need for a cautious approach when interpreting results to ensure accuracy and reliability.

CHAPTER 3: METHODOLOGY

3.1 DATA SOURCES & FEATURE ENGINEERING

TWITTER DATA

In this study, data was gathered from various sources, focusing on three critical types of information that form the foundation of our sentiment analysis: Twitter User Attributes, Tweet Metrics, and Sentiment and Emotional Indices. These data categories were merged into a single dataset to allow for a comprehensive analysis of Twitter sentiment over time. The exact attributes for each category can be seen below:

- **User Attributes:** Twitter User ID, Country, State, City, Gender
- **Tweet Metrics:** Tweet ID, Date of Tweet, Retweet Number, Favorites Number, Mentions Number, Hashtags Number, URLs Number
- **Sentiment and Emotional Indices:** Stress, Fear, Joy, Sadness, Anger, Optimism, and overall Sentiment Scores

TWITTER USER ATTRIBUTES

Twitter user attributes were sourced from Crunchbase, which included the Twitter user ID, country, state, city, and gender of each user. The Twitter User ID helps uniquely identify each user, allowing us to track tweets and analyze patterns per individual. Country, state, and city data enable us to examine how sentiments vary regionally and locally, offering insights into geographic trends in emotional expressions. Lastly, gender data allows for an analysis of how sentiments might differ between genders. This demographic information was essential for understanding the diverse and dynamic ways users interact and express themselves on Twitter.

It is important to note that the analysis was conducted on this a specific set of users identified through Crunchbase, rather than on random Twitter users from the general public. This ensures that the insights and trends identified are relevant to the business and professional communities represented within Crunchbase, offering a targeted understanding of how these users interact and express sentiments on Twitter.

TWEET METRICS

We utilized the Twitter API to collect a substantial dataset encompassing tweets from January 2018 to February 2023. This dataset includes 67,457,434 tweets from 107,613 users across 59 countries, providing a rich basis for our analysis.

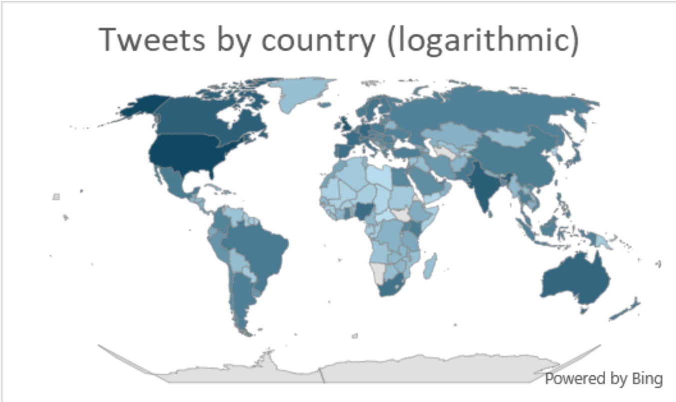
Extracted metrics include the TweetID, which uniquely identifies each tweet, and the timestamp (date), which records when each tweet was posted. Additionally, we gathered metrics such as the Retweet count and Favorites count, which indicate the popularity and user engagement with each tweet. Moreover, the Mentions count was noted as it shows how often other users are tagged within tweets. The Hashtags count and URLs count in tweets were also collected. These metrics are essential for understanding user interactions and their significance on Twitter.

By integrating these Tweet metrics with the Twitter user location, we were able to generate a visual representations that display the geographical distribution of tweets within the dataset.

The highest volumes of tweets originated from predominantly English-speaking countries. This observation is consistent with the limitations of our sentiment analysis, which was designed to assess emotion in English-language tweets. Consequently, the

United States, due to having the largest English speaking population leads with over 40 million tweets, followed by other English-speaking nations such as the United Kingdom, Canada, and Australia. While countries like India and Nigeria also feature prominently on the list, it is likely due to the prevalence of English as in these regions. This linguistic factor has inherently influenced the dataset, steering the majority of analyzable tweets to come from countries where English is widely spoken and used on social media.

COUNTRY	Count
UNITED STATES	40995567
UNITED KINGDOM	7433638
INDIA	3154006
CANADA	2685104
AUSTRALIA	1279445
GERMANY	1185018
FRANCE	656305
NETHERLANDS	580109
NIGERIA	574320
REPUBLIC OF IRELAND	468000
SPAIN	448460



SENTIMENT AND EMOTIONAL INDICES

Several models were used to compute distinct emotional and sentiment indices from tweets, allowing us to conduct a deep analysis of sentiment expression on Twitter.

1.Joy, Optimism, Anger, Sadness, Offensiveness and Hate Detection

For the emotions of joy, optimism, anger, sadness, offensiveness and hate the Twitter-RoBERTa-base model was utilized and fine-tuned on the TweetEval benchmark.

TweetEval is a comprehensive framework designed for Twitter-specific tasks, which assesses emotions using a dataset initially part of SemEval-2018 Task 1. This model evaluates tweets and assigns an emotion index for each of the four emotions on a continuous scale from 0 to 1, where 1 indicates the full presence of the emotion (Barbieri et al., 2020).

2. Fear Detection

Similarly, a RoBERTa-based model was adapted to detect the emotion of fear. This model was retrained on a corpus labeled for emotional content, allowing it to learn context-specific representations of fear. This fine-tuning enhances the model's ability to accurately recognize instances of fear. It scores from 0 to 1, where 1 signifies complete representation of fear in the Tweet (Liu et al., 2019).

3. Stress Classification

The MedDL tool was adapted for detecting stress-related expressions. This approach is based on deep learning techniques with contextual embeddings, providing a sophisticated means to assess stress levels. It uses binary classification to identify stress-related expressions, assigning a score of 0 or 1. A score of 1 indicates the presence of stress, while 0 indicates its absence, enabling monitoring of stress levels expressed within tweets.

4. Sentiment Detection

AllenNLP, a robust NLP library, was employed to analyze overall sentiment levels in tweets. AllenNLP's flexibility in handling the informal and concise language of Twitter, including emojis and hashtags, allows for effective sentiment classification. It categorizes sentiment on a scale from -1 to 1, where -1 represents a completely negative sentiment, 0 neutral, and 1 a fully positive sentiment. (Liu et al., 2019).

COVID-19 DATA

In order to analyze the pandemic's impact on behavior and sentiment, we leveraged data from Our World in Data, a prominent resource known for its comprehensive compilation of global COVID-19 statistics and research. The time frame for the used data spans from

January 2020 to February 2023, capturing the pandemic's onset, its progression, and the subsequent response measures. The data collected provides information about the epidemiological situation and imposed restrictions for each country on daily basis. This data is of great importance as it provides a backdrop against which the public's sentiment and reactions can be compared to.

EPIDEMIOLOGICAL DATA USED:

- **Code (Country):** The ISO country codes are used to uniquely identify each nation in the dataset.
- **Date:** The specific date on which data or policy was recorded.
- **Total Cases:** Cumulative confirmed COVID-19 cases.
- **New Cases:** The number of new COVID-19 cases each day.
- **Total Deaths:** The total number of deaths attributed to COVID-19, reflecting the severity of the pandemic's impact.
- **New Deaths:** Newly reported deaths due to COVID-19.
- **ICU Patients:** The number of COVID-19 patients in intensive care units, shedding light on the strain on healthcare systems.
- **Hospital Patients:** Hospitalized COVID-19 patients.
- **Total Tests:** The aggregate number of COVID-19 tests conducted.
- **New Tests:** Counts of COVID-19 tests conducted in the last 24h.
- **Positive Rate:** The proportion of positive test results, which can indicate the level of viral circulation in the community.
- **Total Vaccinations:** The total count of COVID-19 vaccination doses administered.

COVID-19 RESTRICTION DATA USED:

- **Code (Country):** As above, ISO country codes.
- **Date:** The specific date on which data or policy was recorded.
- **Income Support:** Measures financial support provided to individuals.

- **Workspace Closures:** Status of workplace regulations imposed.
- **Facial Coverings:** Regulations on mask usage.
- **Vaccine Availability:** Accessibility of COVID-19 vaccines.
- **School Closures:** Operational status of educational institutions.
- **Stay Home Requirements:** Mandates for populations to remain at home, influencing social mobility and contact rates.
- **Close Public Transport:** Suspension of public transport services.
- **International Travel Controls:** Restrictions on international travel, aimed at controlling the spread across borders.
- **Restriction Gatherings:** Limits on group sizes for gatherings.
- **Debt Relief:** Economic policies aimed at alleviating financial burdens during the pandemic.
- **Vaccination Policy:** Strategies and prioritization schemes guiding vaccine rollout.
- **Containment Index:** A composite measure of the stringency of policy responses to COVID-19.

Understanding these details is really important because they show us how the COVID-19 pandemic has been unfolding and how well different countries' strategies are working. We want to examine if these factors have an impact on what people feel and say about the pandemic on social media.

TWEET TOPIC MODELING USING ROBERTA

ROBERTA MODEL

For the purpose of identifying and categorizing the thematic content of tweets, this study employed the RoBERTa (Robustly Optimized BERT Pretraining Approach) model, a variant of the BERT architecture known for its enhanced training methodology and performance on language understanding tasks (Liu et al., 2019). RoBERTa is able to capture contextual meanings and nuances, making it particularly suitable for analyzing the informal and diverse language found in tweets.

DATASET AND PREPROCESSING

The RoBERTa model used in this study was fine-tuned on a subset of the New York Times headlines dataset, publicly available on Kaggle, consisting of 256,000 article titles published from 2000 to 2023.

Prior to any further topic-related analysis, the dataset was filtered so that only entries with topic confidence score greater than 0.9 were retained, ensuring that the training data consisted of highly relevant and accurately labeled instances.

TRAINING PROCEDURE

The fine-tuning of the RoBERTa model was conducted with the following hyperparameters:

- learning rate: 5e-05
- training batch size: 8
- evaluation batch size: 8

The model was trained over 5 epochs, with a linear learning rate scheduler and a warm-up step of 500 to prevent overfitting at the initial stage. Adam optimizer was used with betas set to (0.9, 0.999) and epsilon at 1e-08, which are standard configurations known to work well with transformer models.

The model was trained to classify tweets into one of the eight predefined categories listed below:

1. Sports
2. Arts Culture and Entertainment
3. Business and Finance
4. Health and Wellness
5. Lifestyle and Fashion
6. Science and Technology

7. Politics
8. Crime

MODEL EVALUATION AND PERFORMANCE

The performance of the topic classifier was evaluated on a test set comprising 51,200 cases. The model achieved an accuracy, precision, recall, and F1 score of 0.91. These metrics indicate a high level of consistency and reliability in the model's classification capabilities.

The performance was further dissected by individual topic categories, which can be found in the table below.

Topic	Precision	Recall	F1	Support
Sports	0.97	0.98	0.97	6400
Arts, Culture, and Entertainment	0.94	0.95	0.94	6400
Business and Finance	0.85	0.84	0.84	6400
Health and Wellness	0.9	0.93	0.91	6400
Lifestyle and Fashion	0.95	0.95	0.95	6400
Science and Technology	0.89	0.83	0.86	6400
Politics	0.93	0.88	0.9	6400
Crime	0.85	0.93	0.89	6400

3.2 ANALYTICAL APPROACH

The analysis performed in this study was strategically structured around three geographical scopes: worldwide, the United States of America (USA), and Sweden. USA was selected due to it having the highest volume of tweets in our dataset, offering a substantial sample for in-depth sentiment analysis and providing a broad perspective on the impact of COVID-19 in a country with significant global influence. On the other hand Sweden was chosen for its unique approach to handling the COVID-19 pandemic, which differed from the lockdown strategies adopted by most countries, thus presenting an intriguing case study. Each geographical scope entailed a series of analytical steps using statistical and machine learning methods, primarily executed in Python and R, as detailed below:

WORLDWIDE ANALYSIS

1.Correlational Studies

We began by exploring correlations among the different emotional indices to understand the relationships between various emotions expressed in tweets. Additionally, minimum, maximum, and average values of each emotion were examined. Furthermore, correlations between COVID-19 data (epidemiological and restriction related) were examined to identify any significant associations.

2. Sentiment Fluctuations Over Time

Sentiment fluctuations on a global scale were graphed. Each emotion (joy, anger, fear, optimism, hate, sadness, offensiveness, anger, and sentiment) was graphed over time to visualize changes and trends throughout the pandemic.

USA & SWEDEN ANALYSIS

1.Event Research

We conducted a detailed review of major political and social events in the USA and Sweden from 2018 to 2023 to provide a contextual backdrop for the sentiment analysis. We also researched governmental responses, public health measures, economic policies implemented by each country in response to the pandemic.

2. Sentiment Fluctuations Over Time

Similar to the worldwide analysis, we graphed the fluctuations of each emotion over time, both in relation to the average of each emotion and as standalone indices. Emotion fluctuations were examined alongside epidemiological data and COVID-19 restriction measures to understand the impact of these factors on public sentiment.

4. Statistical Testing

T-tests were performed to compare emotional metrics before and after the onset of the pandemic. Additionally, we utilized fixed effects regression models, investigating the impact of our emotional metrics (joy, fear, hate, etc.) as dependent variables. Tweet characteristics (favorites, retweets, etc.) were included as control variables. An instrumental variable (IV) was also introduced to denote tweets during the COVID-19 period, assigned a value of 1, while those before the pandemic were assigned a value of 0.

TOPIC MODELING

1.Categorization

Topic modeling techniques as described above were applied to categorize tweets into one of nine predetermined categories. This method allowed us to group tweets by similar subject matter, making it possible to examine patterns and trends within specific topics of discussion.

2.Topic Popularity and Sentiment Changes

Once the tweets were categorized, we conducted a thorough analysis to track the popularity of each topic over time on a worldwide scale. We measured how frequently each topic was discussed in tweets during different periods of the study. Alongside popularity, we also analyzed how the sentiments expressed in these tweets evolved. By examining sentiment changes within each topic category, we could identify shifts in public mood or opinion related to specific subjects.

3.3 CHALLENGES & LIMITATIONS

This study, while comprehensive in its approach and methodology, is subject to several limitations that may affect the interpretation of the findings. Understanding these limitations is crucial for placing the results in the proper context.

LANGUAGE AND CULTURAL BIAS

The sentiment analysis was conducted on English-language tweets. This introduces a language bias, limiting the analysis to predominantly English-speaking users and potentially overlooking the sentiments expressed in other languages. Additionally, cultural nuances in language usage and expression of emotions may not be fully captured due to the limited fluency of some users whose first language isn't English, which can affect the accuracy of sentiment and emotion detection models. This limitation restricts the analysis to only those users in non-English speaking countries who tweet in English, leaving the sentiments and opinions of other language speakers unrepresented. Such constraints could lead to an incomplete picture of the Twitter landscape, underrepresenting the diverse voices and potentially skewing the insights derived from the data toward English-speaking perspectives. This aspect is crucial to consider, as it affects the generalizability and applicability of the findings across different linguistic and cultural groups.

SOCIAL MEDIA BIAS

The data collected via the Twitter API and other sources like Crunchbase may not comprehensively represent the global Twitter population. Factors such as selection bias of users who choose to make their tweets public could skew the dataset. Such constraints might lead to an incomplete picture of the Twitter landscape, particularly affecting the representation of less active or private users.

Additionally, Twitter users might not represent the general population accurately, skewing the analysis towards younger individuals. This demographic bias may not accurately reflect the sentiments of older or less digitally engaged populations, potentially leading to a distorted view of public sentiment.

MODEL DEPENDENCY

The reliance on models such as RoBERTa, AllenNLP, and MedDL for sentiment analysis and topic modeling may also pose limitations. While these models are state-of-the-art, they are not infallible and can be prone to errors, especially in the context of complex and informal text like tweets. The performance of these models is heavily dependent on the training data, and any biases or shortcomings in these datasets could propagate to the analysis.

TEMPORAL COVERAGE

This study covers a substantial period from January 2018 to February 2023, where the dynamic nature of social media and public sentiment means that findings may not be applicable outside of this timeframe. The period analyzed was profoundly influenced by numerous global events, including the COVID-19 pandemic, which co-occurred with other significant social, economic, and political changes. These factors intricately shaped public discourse and sentiment on Twitter. Consequently, the results of the sentiment analysis are dependent on the situational context of this time period. Any interpretation of the data must consider these influences, as they are inseparable from the broader landscape of the time.

IMPACT OF MAJOR EVENTS

The study period includes several major global events beyond COVID-19, such as the George Floyd protests, the 2020 US presidential election, and the war in Ukraine. These events significantly influenced public sentiment and may confound the analysis of pandemic-specific emotional shifts. Detaching the effects of these overlapping events poses a significant analytical challenge.

INCOMPLETE UTILIZATION OF TWEET INFORMATION

The analysis relies solely on the textual content of tweets, overlooking other forms of information such as images and videos. These elements can convey significant emotional and contextual information that is not captured through text analysis alone.

COUNTRY-LEVEL ANALYSIS LIMITATIONS

The analyses for the USA and Sweden are conducted at the country level, ignoring regional variations in sentiment. This approach fails to account for the differences within these countries, especially in the USA where states implemented varying measures in response to the pandemic. Ignoring these intra-country variations may oversimplify the analysis and obscure important local differences in public sentiment.

ETHICAL CONSIDERATIONS

Analyzing user-generated content on social media raises ethical concerns, particularly around privacy and consent. Users may not be aware that their tweets are being analyzed, leading to potential ethical dilemmas. The aggregation and interpretation of data must be handled with care to avoid misrepresenting individuals or groups.

CHAPTER 4: RESULTS

4.1 WORLDWIDE ANALYSIS

CORRELATION STUDIES

In this section, we delve into two main types of correlations to understand the interconnections between different aspects of the data. First, we explore correlations between different emotions expressed in tweets to identify patterns and relationships in how sentiments are conveyed during the pandemic. This analysis helps reveal how emotions like joy, anger, and sadness interact and influence each other amidst a global crisis. Second, we examine correlations between COVID-19 data, which includes both epidemiological metrics (such as case counts and deaths) and government-imposed restrictions (like lockdowns and travel bans). This exploration aims to understand how these external factors correlate with changes in public sentiment.

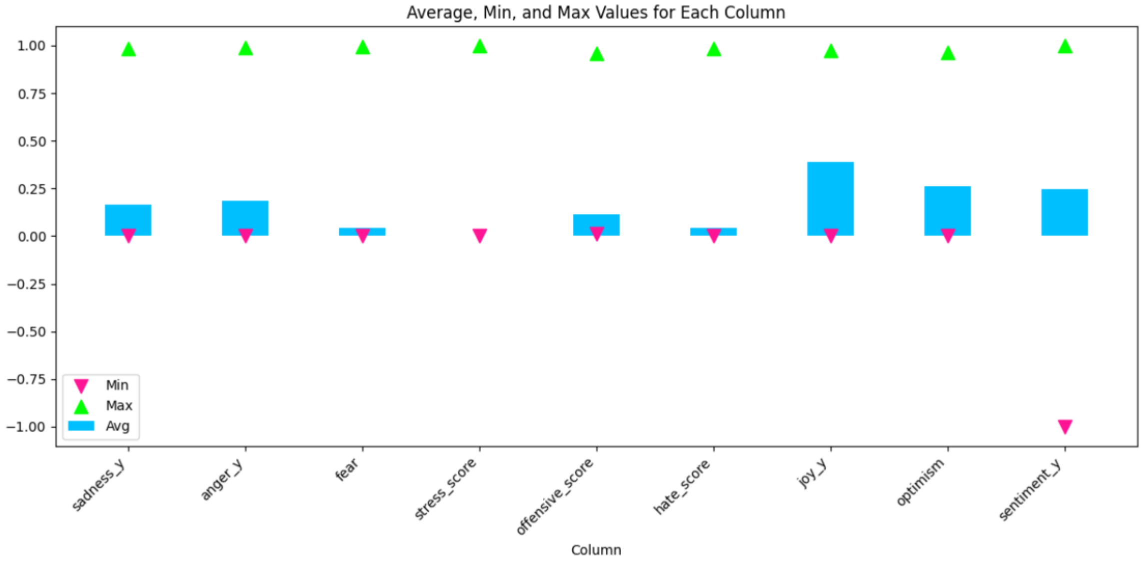
OVERVIEW OF EMOTIONAL INDEXES

In examining the landscape of Twitter during the COVID-19 pandemic, our analysis extends to the minimum, maximum, and average values recorded for each emotion.

Emotion	Min	Max	Avg
sadness_y	0.003	0.987	0.164
anger_y	0.003	0.988	0.186
fear	0	0.997	0.041
stress_score	0	1	0.001
offensive_score	0.014	0.957	0.114
hate_score	0.003	0.986	0.044
joy_y	0.002	0.976	0.387
optimism	0.002	0.964	0.263
sentiment_y	-1	1	0.248

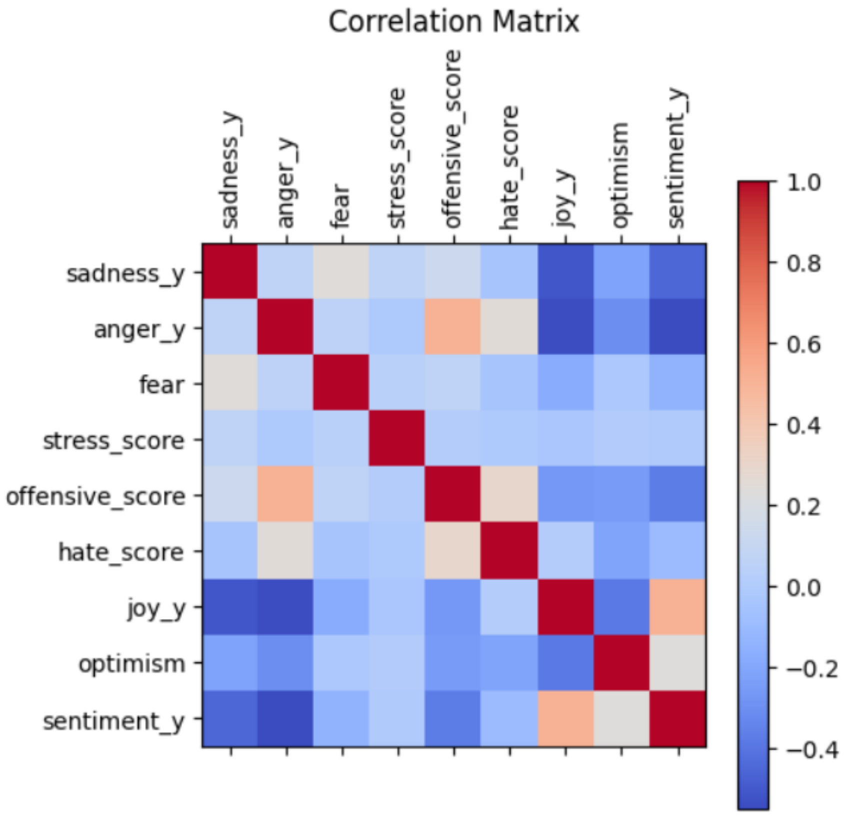
It can be observed that all emotions cover the full span of values of their respective indices, suggesting that a wide variety of emotions were present in our dataset. Looking at the bigger picture, it seems that on average, positive emotions like joy and optimism have higher average indices than the negative ones like sadness or anger.

It is worth noting that the average score for stress is notably low at 0.001, which implies that expressions of stress were infrequently captured in our dataset. This could be due to a lower prevalence of stress expressions on Twitter or a limitation of our model in the detection and measurement of stress. Given this low average, the fact that stress did not show significant correlations with other emotions, and the chaotic shifts in stress we later observed, we have opted not to pursue further analysis on the change of stress within the context of this study.



CORRELATIONS BETWEEN EMOTIONS

The correlation matrix provided is a visual representation of how different emotions expressed in tweets relate to each other during the COVID-19 pandemic.



Positive correlations:

Offensive Score and Anger: The dataset exhibits a strong positive correlation between offensive score and anger, which aligns with conventional expectations. This correlation shows the frequent co-occurrence of offensive language with the expression of anger in tweets.

Sentiment and Joy: A positive correlation is observed between overall sentiment and the expression of joy. This correlation indicates that tweets with a higher sentiment score to also contain expressions of joy.

Negative correlations:

Joy and Anger/Sadness: A significant negative correlation is observed between joy and both anger and sadness. This finding supports the contrasting nature of these sentiments.

Joy and Optimism: The correlation matrix also reveals a somewhat unexpected negative correlation between joy and optimism. This relationship suggests that as joy is expressed more frequently in tweets, optimism tends to be less prevalent, and vice versa. It might indicate that while joy represents an immediate emotional response to specific events or moments, optimism may capture a broader, forward-looking perspective.

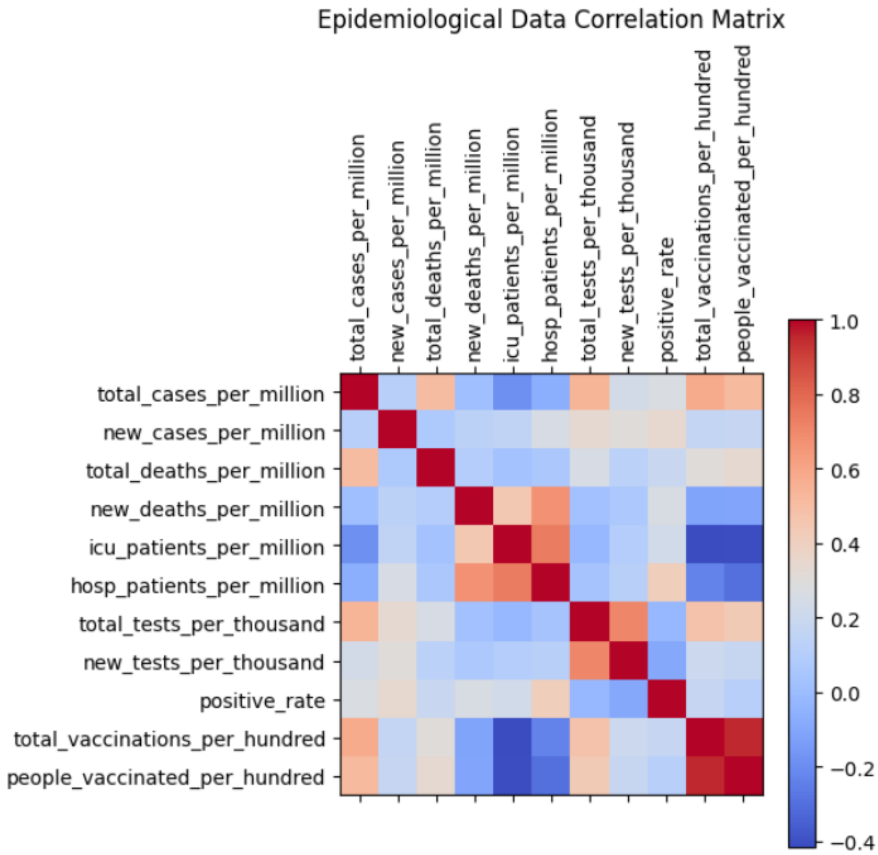
Sentiment and Anger/Sadness/Offensiveness: The relationship between sentiment and these three emotions reveals that they are negatively correlated. Anger has the strongest inverse association with sentiment, indicating that tweets expressing anger tend to correspond with the lowest sentiment scores. This is followed by sadness and offensiveness, which also negatively affect sentiment but to a slightly lesser degree.

Interestingly, the emotion of stress appears to show no substantial correlation with the other emotions. This suggests that the expression of stress on Twitter might operate independently of the other emotions captured by our dataset or may be influenced by factors not directly measured by the other metrics. It's important to note that stress was measured using a different tool that assigned binary values to the tweets, which could account for the lack of observed correlation. Given this low average, and the fact that stress did not show significant correlations with other emotions, we have opted not to pursue further analysis of the change in stress within the context of this study.

CORRELATIONS BETWEEN COVID-19 DATA

In the subsequent section, we shift towards the COVID-19 data, dissecting correlations within two categories: epidemiological metrics and restriction measures.

CORRELATIONS BETWEEN EPIDEMIOLOGICAL DATA



Positive Correlations:

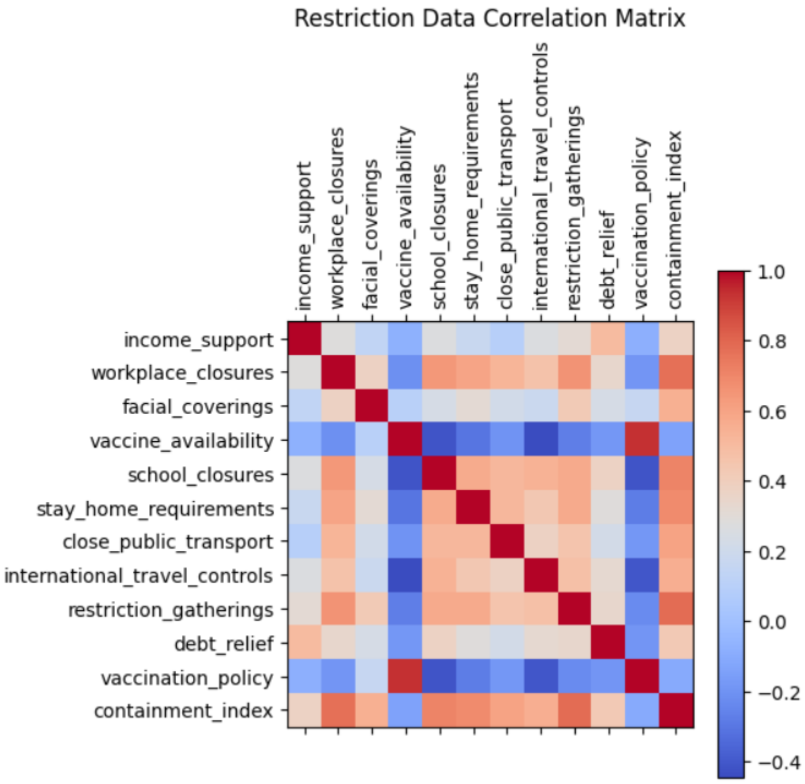
The highest positive correlation index, nearing almost 1, can be found between the indices of the total vaccinations per hundred and people vaccinated per hundred. For this reason, they shall be encapsulated under the term “vaccinations” from now on.

Another high correlation can be observed between the total cases and total deaths per million, which is expected due to them being closely intertwined; as the cases increase, so too does the mortality rate.

Furthermore, the correlation between new deaths and hospital patients per million displays the link between serious COVID-19 cases and subsequent deaths. Similarly, the relationship between ICU patients and hospital patients per million is equally as strong.

Negative Correlations:

The inverse relationship between vaccinations and hospital patients and ICU patients is quite strong, indicating a clear pattern: as more people get vaccinated, there's a decrease in severe cases that require hospitalization. Building upon the influence of vaccinations, the relationship between vaccinations and new deaths per million is similarly negative. These connections suggest that vaccines not only helped prevent infection but also reduced the severity and mortality of COVID-19.



Positive Correlations:

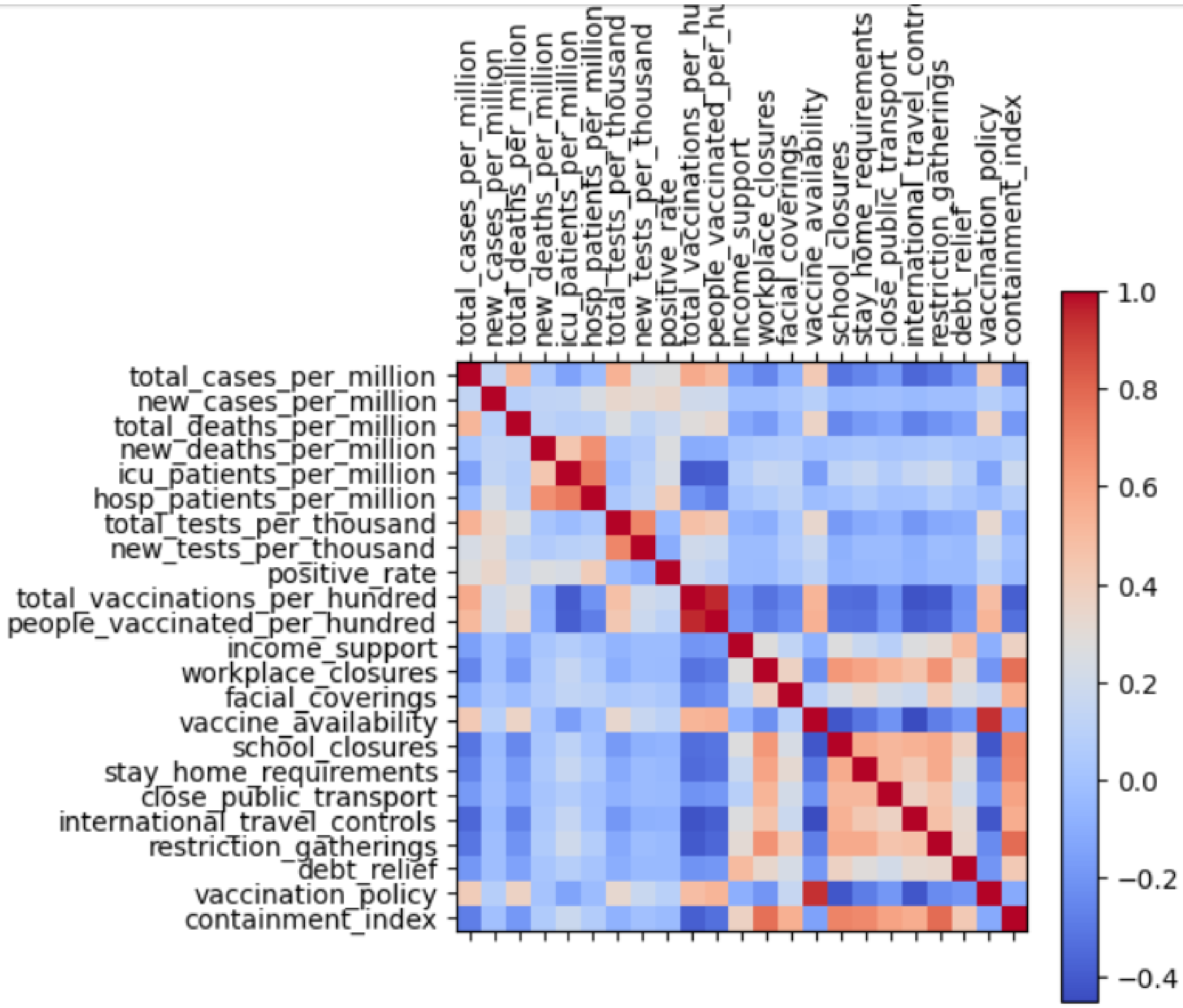
A detailed analysis of how different COVID-19 restriction measures are connected shows that, workplace closures, school closures, stay-home requirements, closure of public transport, international travel restrictions and restrictions on gatherings not only individually aligned with the increase of the containment index but are also very closely correlated with each other. This connection implies that these measures are often implemented together.

Interestingly, mandates on facial coverings show a very slight correlation with the aforementioned cluster of restrictive measures, but have a significantly higher correlation with the containment index. This observation suggests that the enforcement of facial coverings operates almost independently of the broader, more invasive restrictions.

Negative Correlations:

On the other hand, there's a noticeable negative connection between vaccine availability, vaccination policies, and the previously mentioned restrictive measures like school closures, stay-at-home orders etc. This shows a clear opposite relationship: as vaccines become more widely available and vaccination policies become more established, there's less need for strict social restrictions.

CORRELATIONS BETWEEN EPIDEMIOLOGICAL DATA & COVID-19 RESTRICTIONS

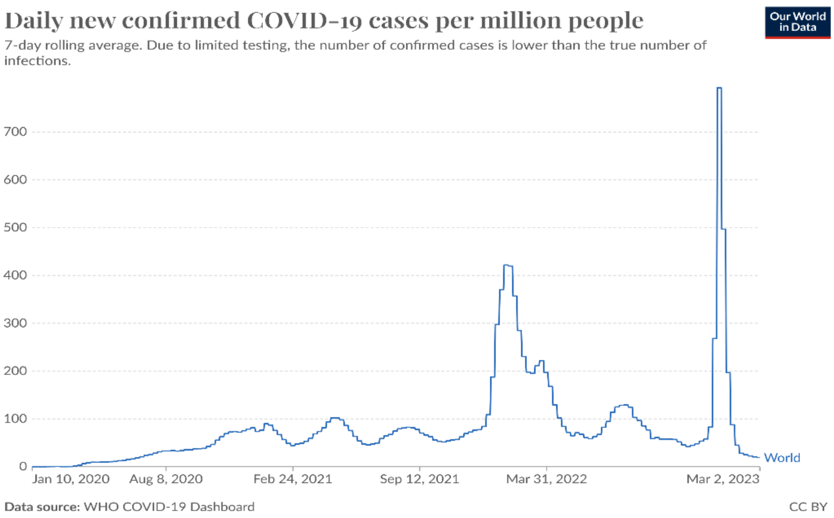


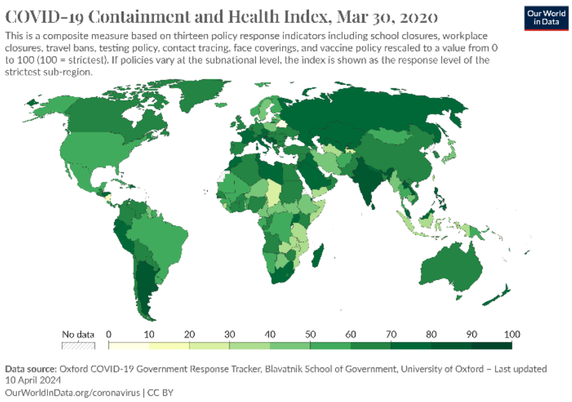
The correlation matrix between epidemiological data and COVID-19 related restrictions offers some interesting insights. The anticipated strong relationships between the severity of the pandemic and the stringency of restrictions are not as evident. This might initially seem counterintuitive since one would expect that as the number of cases and deaths rise, so would the level of restrictions.

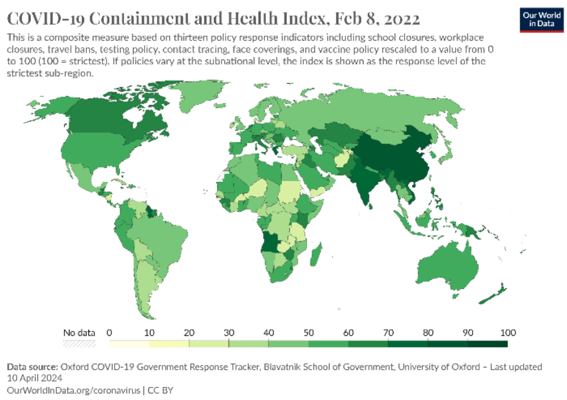
This could be explained by the fact that the implementation of strict measures occurred early in the pandemic, after the initial outbreak where the absolute numbers of cases and deaths were still comparatively low. Consequently, during later waves, although epidemiological figures were significantly higher, the restrictions were not increased proportionately, possibly due to a variety of reasons, including policy fatigue, economic pressures, or a better understanding of targeted measures.

The data obtained from "Our World in Data" fortifies the hypothesis presented. This idea is supported by comparing the containment and health index at two different times.

Early on in the pandemic, (March 30th, 2020 data used), stringent measures were implemented worldwide despite the reported cases being relatively low (6.3 cases per million in this case). The containment and health policies were in their most aggressive phase, reflecting a global stance of extreme caution in response to the health crisis. In contrast, during later COVID-19 waves, (February 8th, 2022 data used), when the epidemiological data pointed to a significantly higher number of cases (357 cases per million), the containment measures depicted by the index were notably more relaxed. This difference supports the observation that the link between policy strictness and epidemiological indicators isn't as strong as expected.







4.2 CASE STUDY: USA

APPROACH TO THE COVID-19 PANDEMIC

The management of the COVID-19 pandemic in the USA has been characterized by a dynamic interaction between health initiatives, political opinions, and public reactions. Partisan divisions and varying levels of trust in government and media sources have influenced the implementation and effectiveness of public health measures. Understanding these dynamics is crucial for understanding the analysis of our data.

Early Phase Response:

During the early phase, the U.S. implemented various health initiatives and public health measures to combat the spread of the virus.

Clinical trials were launched to assess potential treatments, indicating a proactive approach to addressing the medical aspects of the pandemic (Mehta et al., 2020). Public health measures such as lockdowns and social distancing were implemented, although their effectiveness varied across states and were influenced by political factors (Milani, 2020). Political opinions played a significant role during this phase, with partisan differences shaping responses to COVID-19. These differences influenced public behavior and state-level policy decisions. (Gadarian, Goodman, & Pepinsky, 2020).

The public response during the early stages of the pandemic reflected a mix of behavioral, emotional, and psychological reactions. Health precautions such as social distancing and mask-wearing were widely adopted, but there was also a notable decline in visits to emergency departments for non-COVID-19 related emergencies, indicating public hesitancy to seek medical care (Hartnett et al., 2020; Lange et al., 2020).

Information-seeking behavior increased following announcements of COVID-19 cases, accompanied by heightened levels of stress and anxiety due to uncertainty about the virus and government responses (Bento et al., 2020).

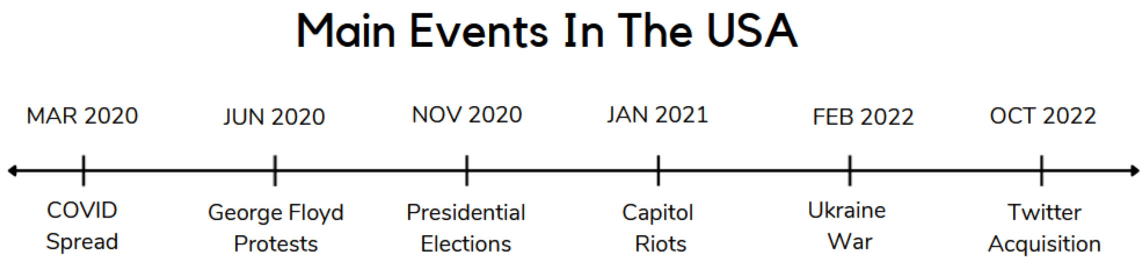
Media consumption and political opinions significantly influenced public behavior and attitudes. Trust in different media outlets and partisan responses shaped compliance with health advisories and perceptions of the pandemic's severity (Gadarian, Goodman, & Pepinsky, 2020).

Late Phase Response (After Vaccines):

The introduction of vaccines marked a critical turning point in the pandemic response in the U.S. Efforts shifted towards vaccine rollout and achieving herd immunity. Public health campaigns were created aimed to promote vaccines, which was a polarizing topic along political parties (Hu et al., 2021).

Despite the availability of vaccines, political opinions continued to shape public responses. Partisan divisions persisted, influencing attitudes towards vaccination and trust in public health information (Gadarian, Goodman, & Pepinsky, 2020).

MAIN EVENTS



June 2020 - George Floyd Protests

The George Floyd protests began in Minneapolis in late May 2020, following the police killing of George Floyd, an African American man. Sparking widespread civil unrest, these protests focused on issues of racial injustice and police brutality. The movement quickly spread across the United States and globally, significantly influencing public discourse and leading to calls for police reform and racial equality. These events also impacted social media sentiment, becoming a major topic of discussion and activism on platforms like Twitter. (Ahaotu, 2023))

November 2020 - Presidential Elections

The 2020 United States presidential election was highly controversial, featuring incumbent President Donald Trump and former Vice President Joe Biden. The election was characterized by a polarized political climate, widespread misinformation, and intense scrutiny over voting processes due to the pandemic. Biden's victory marked a significant political shift and was heavily discussed on social media. (Hopkins, D. J., 2021)

January 2021 - Capitol Riots

On January 6, 2021, supporters of Donald Trump stormed the U.S. Capitol in an attempt to overturn his defeat in the 2020 presidential election. This event marked a significant moment in U.S. history, highlighting deep political divisions and leading to debates over democracy and security in the Capitol. The riots were a major point of discussion on social media, impacting public sentiment and leading to widespread condemnation. (Goldsmith, J., 2021)

February 2022 - Ukraine War

While not an event within the U.S., the invasion of Ukraine by Russia in February 2022 had global implications, including significant political and economic impacts in the U.S. It led to a surge in energy prices and prompted the U.S. to reevaluate its geopolitical strategies and foreign policy. The event significantly affected American public sentiment, particularly regarding international relations and national security. (Goldthau, A., & Hughes, L., 2022)

October 2022 - Twitter Acquisition

Elon Musk's acquisition of Twitter in October 2022 was a significant event in the tech and media landscape. The takeover by one of the world's richest individuals was met with mixed reactions and raised questions about free speech, misinformation, and the role of social media in public discourse. (Pah, H., 2023)

Following the acquisition, Elon Musk implemented several notable changes:

- **Leadership and Staff Changes:** Significant layoffs and restructuring within the company, including the dismissal of top executives.

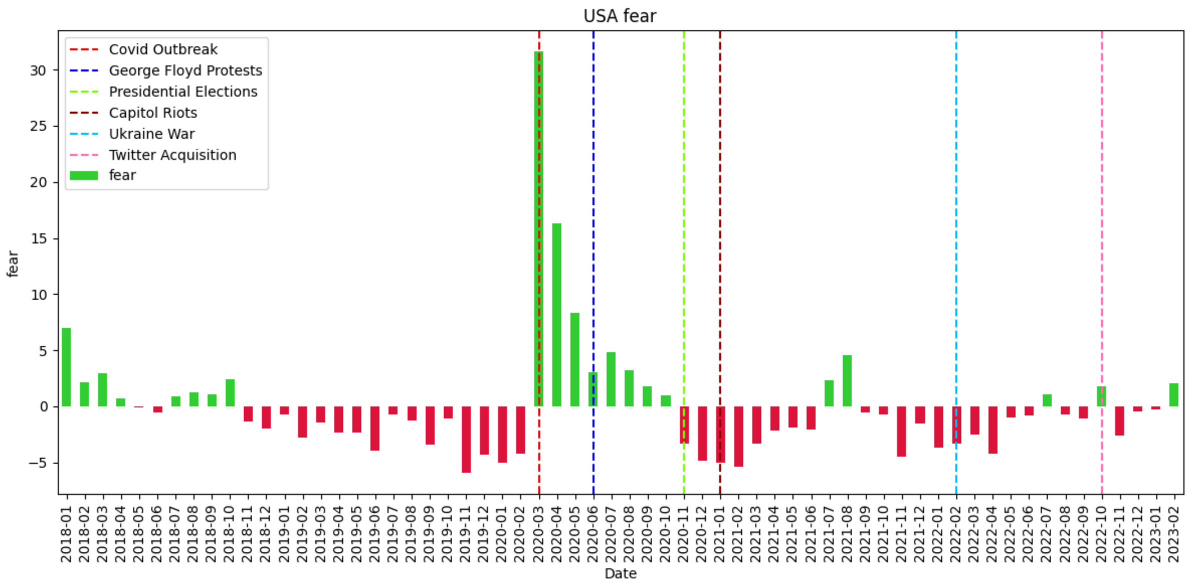
- **Content Moderation Policies:** Adjustments to content moderation policies, aiming to promote free speech while balancing the need to manage misinformation and harmful content.
- **Algorithm and Feature Tweaks: Modifications** to the platform's algorithm to prioritize tweets from paying subscribers and other changes aimed at improving user engagement.

These changes have fundamentally reshaped how users interact on the platform, influencing the nature of Tweets, user engagement, and overall user experience.

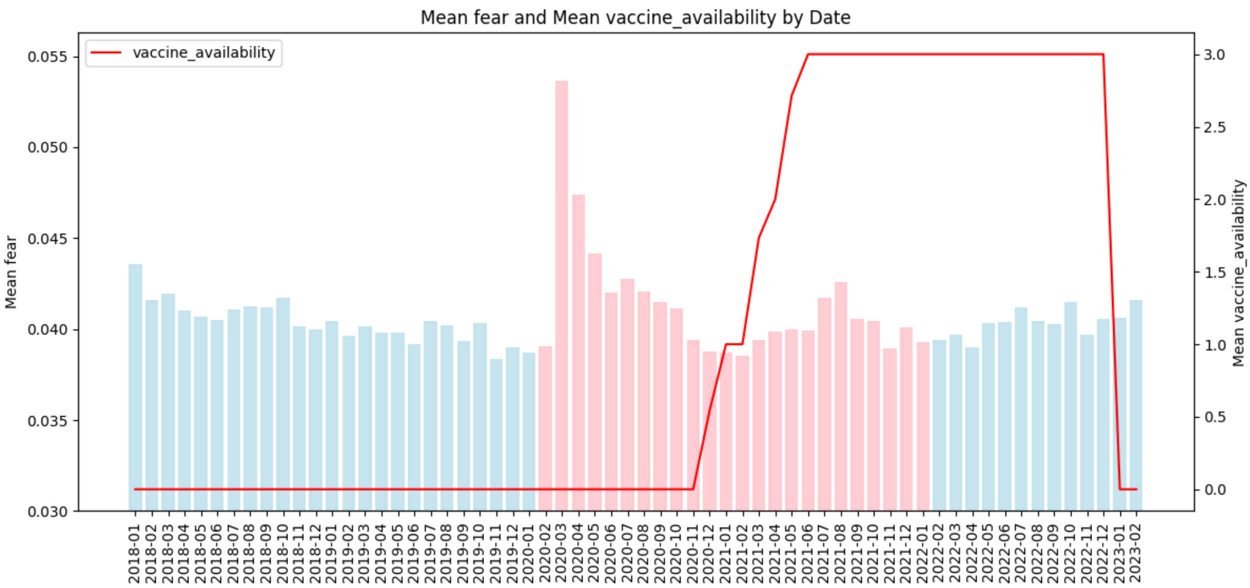
SENTIMENT FLUCTUATIONS

The following graphs offer an analytical perspective on the emotional landscape of the United States, as seen through the lens of Twitter data. These display the percentage change in average monthly emotion, juxtaposed against the mean emotional index of the country. By tracking those changes in emotional states, we can not only show when emotions are stable but also when they significantly differ from the usual pattern, indicating how the nation reacts to different events.

FEAR

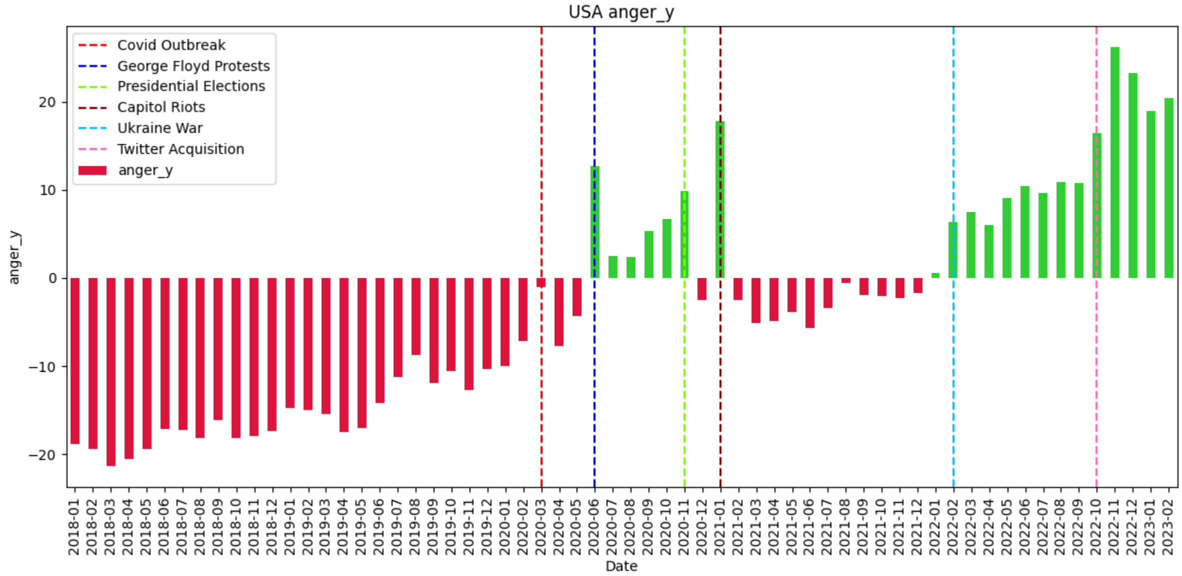


This graph showcases the fluctuations in fear in tweets, depicting a notable spike in the levels of fear during the initial phase of the COVID-19 pandemic. The highest peak occurs during March 2020, when COVID-19 was spreading throughout the US and the globe. In the months following the initial outbreak, a gradual stabilization is observed, possibly reflecting the population's adaptation to the new circumstances imposed by the pandemic.



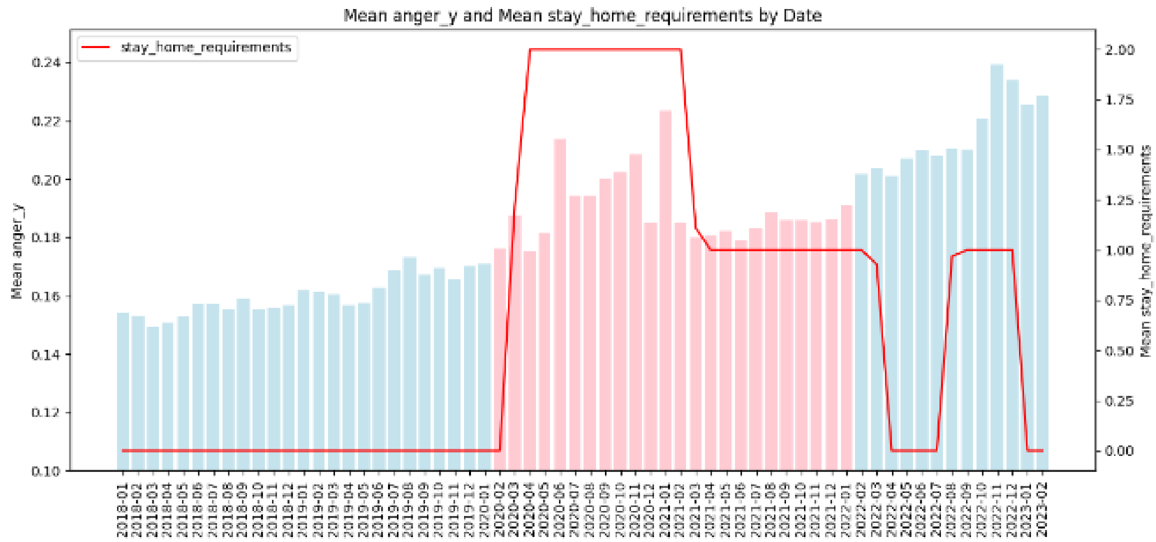
The above graph presents the average monthly fear index and vaccine availability in the United States. We can see a clear relationship between fear levels and vaccine availability in the United States, as the red line indicating vaccine availability rises following the vaccine rollout announcement in December 2020, there is a significant drop in the average monthly fear. This trend suggests that the availability of vaccines played a crucial role in reducing public anxiety.

ANGER



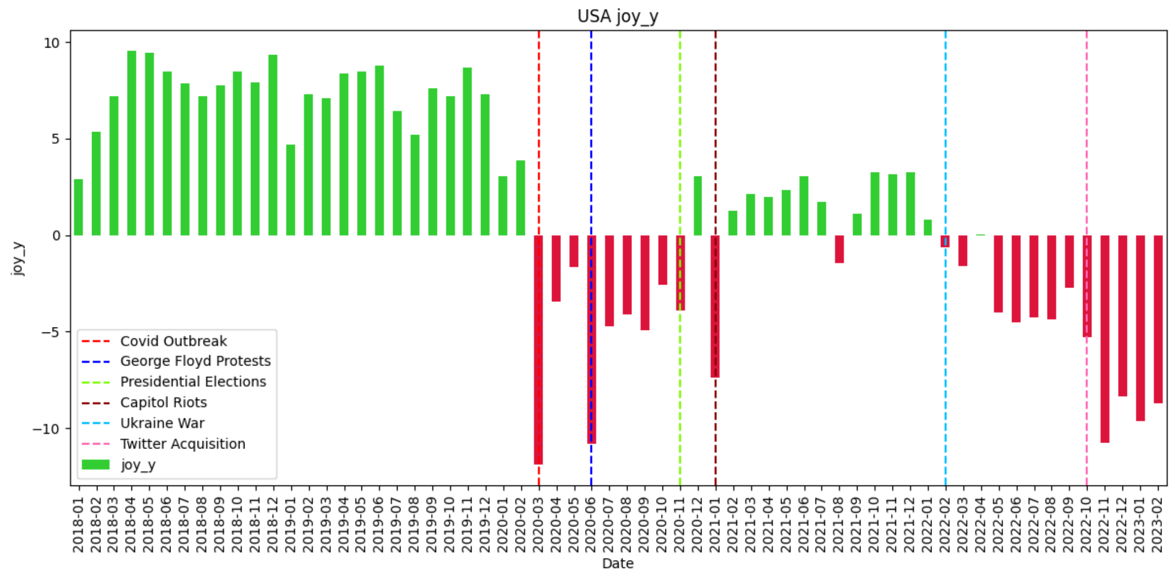
This graph portrays the fluctuations in anger within the United States. From early 2018 to the pandemic's start, there's a steady growth in anger. A significant rise in March 2020 coincides with the emerging COVID-19 crisis. The notable spike in June 2020 aligns with the protests sparked by George Floyd's death, highlighting societal unrest. The 2020 presidential election and subsequent January 2021 Capitol events further escalate anger levels at those moments in time.

Subsequently, anger returns to below average levels, before skyrocketing following Russia's invasion of Ukraine. Anger levels remain high and escalate even further following the acquisition of Twitter by Elon Musk. Due to the less limiting content moderation policies implemented under Musk's leadership, there was more allowance for angry tweets, and they were less suppressed by the algorithm. The fact that the highest peak corresponds with the acquisition of Twitter by Elon Musk is not surprising, given that it was highly unpopular amongst Twitter users and led to an environment where negative sentiments could be more freely expressed. The same is true for other negative emotions.

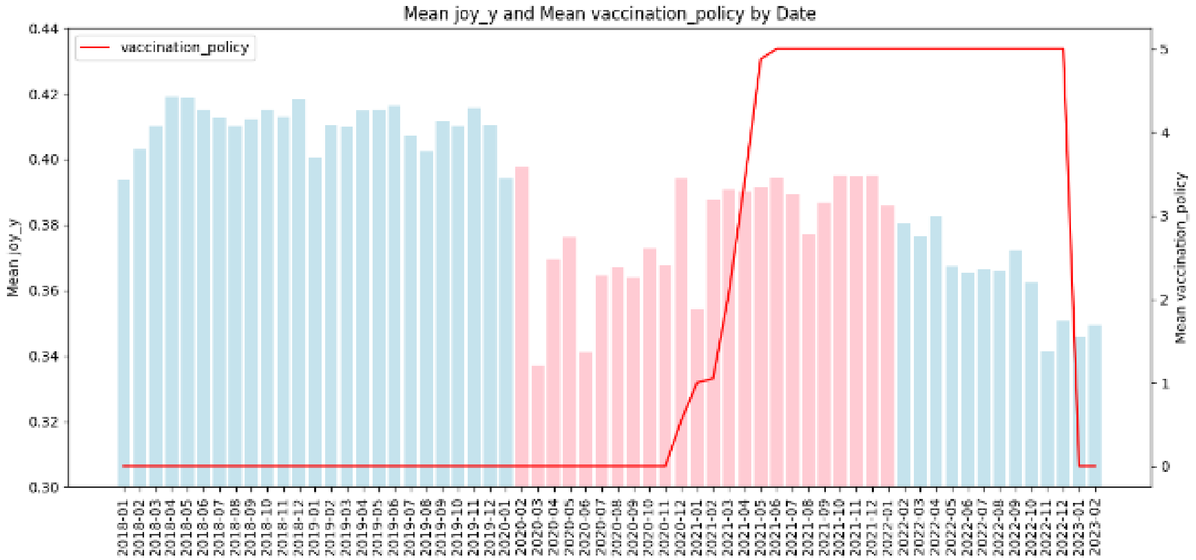


When comparing the anger levels to the stay-at-home requirements we can see that anger remains elevated during the COVID-19 period with levels significantly dropping when restrictions are loosened around March of 2021. However, we should not forget that the heightened anger during that time is also due to the intersection with other significant events as mentioned above.

JOY

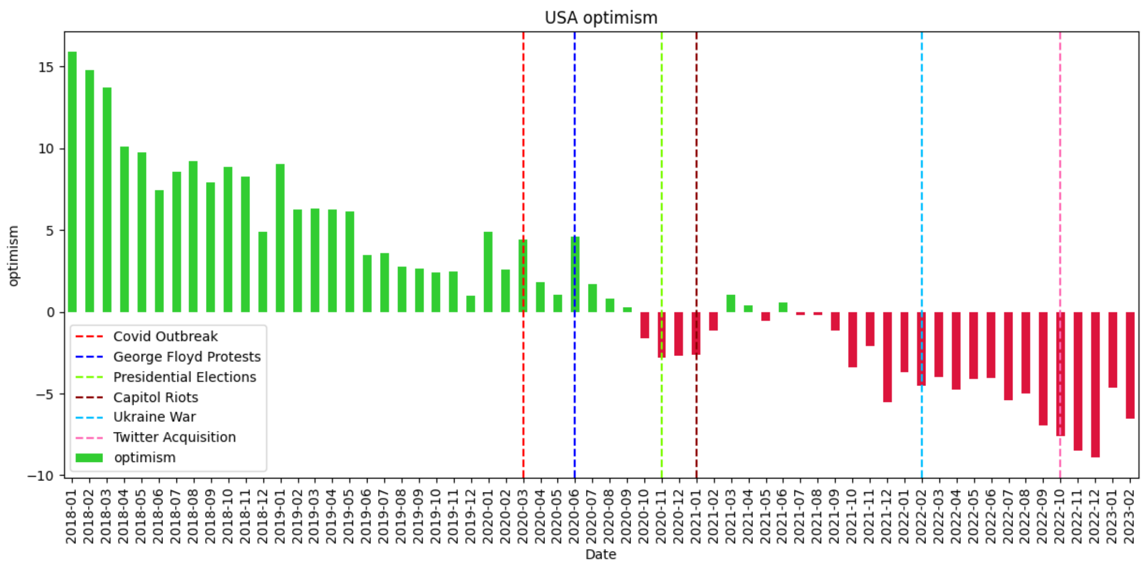


Joy, in many ways behaves opposite to anger. At the start of the COVID-19 pandemic, there's a sharp decrease in happiness, which remains consistently low throughout the rest of 2020, partially influenced by the political and social events of the time. This shows how deeply the pandemic has affected the overall well-being of society.



As vaccine rollouts commence in December 2021, a resurgence of joy is evident, signaling a return to optimism. This recovery, however, is short-lived as the escalation of the conflict in Ukraine in early 2022 plunges joy back into negative values. In October 2022, coinciding with Elon Musk's acquisition of Twitter, the joy index plummets dramatically, almost mirroring the initial COVID-19 response.

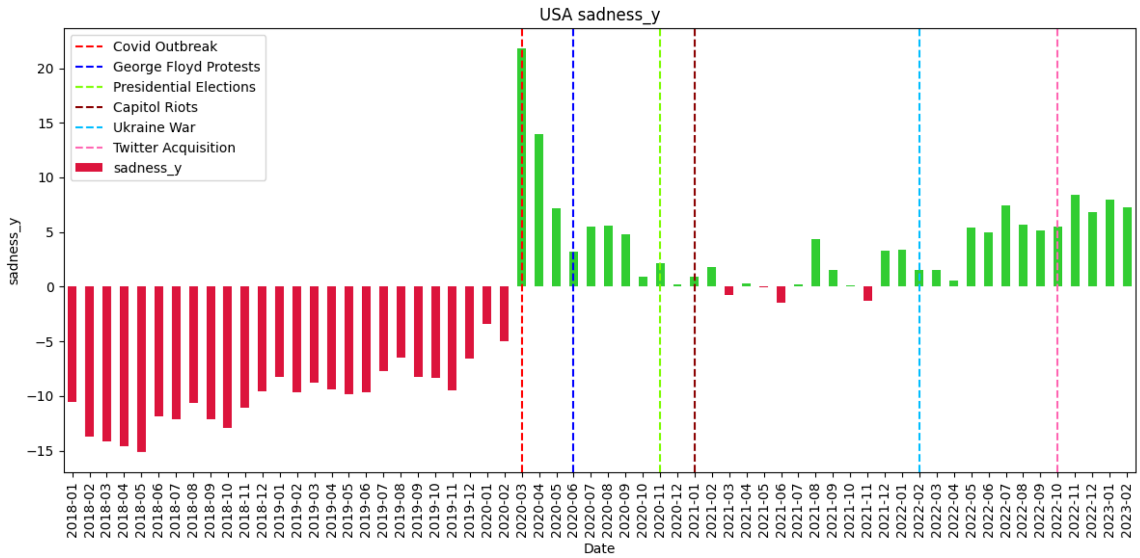
OPTIMISM



The graph depicting the monthly percentage change in optimism demonstrates a distinct pattern compared to other emotional indicators. Initially showing resilience at the onset of COVID-19, optimism doesn't exhibit a sharp difference as seen with emotions such as fear or joy.

The gradual decline in optimism over time, as shown in the graph, may be attributed to the cumulative effect of prolonged exposure to negative news and the constant presence of challenging global events on social media platforms like Twitter. Research by Oh, Goh, and Phan (2018) suggests that such exposure can significantly influence digital news consumption and sharing behaviors, leading to a shift in collective sentiment towards a more cautious or even pessimistic outlook on platforms like Twitter.

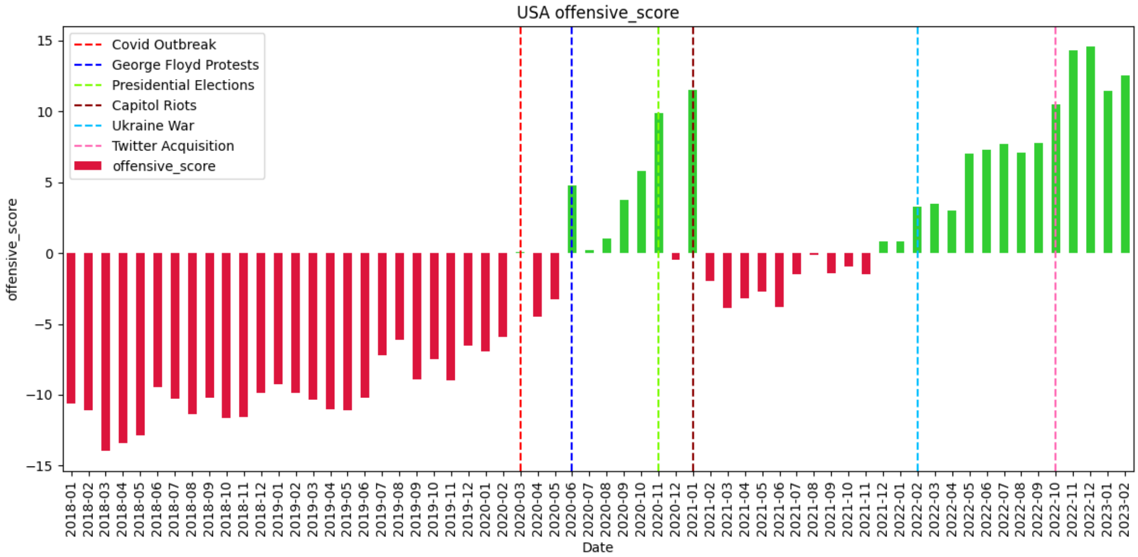
SADNESS



The graph highlights the fluctuations in expressions of sadness on Twitter over time, pinpointing two significant surges that align with major world events. In March 2020, there's a pronounced spike that correlates with the initial spread COVID-19 and the imposition of lockdowns in the USA.

A second, more gradual increase in sadness is observed in 2022, starting with the commencement of Russia's invasion of Ukraine. This period sees an upswing in sadness, perhaps as a reaction to the widespread humanitarian distress and the escalation of conflict.

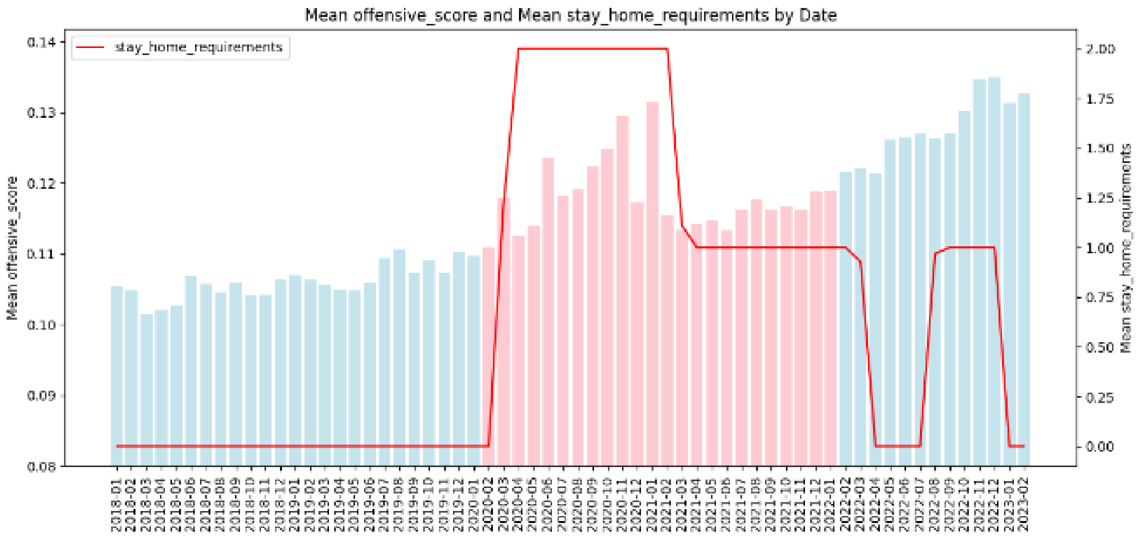
OFFENSIVENESS



The graph tracing the monthly percentage change in offensiveness on Twitter indicates a general uptrend over time even before the onset of the pandemic, suggesting an increase in the use of offensive language on the platform. This can be attributed to the same causes as the gradual decrease in optimism over time (Oh, Goh, and Phan, 2018).

Notably, there are spikes in offensiveness leading up to the 2020 U.S. presidential elections and a significant peak around January 2021, which aligns with the Capitol attack. These events likely fueled polarizing discussions, reflected in the increase in offensive expressions.

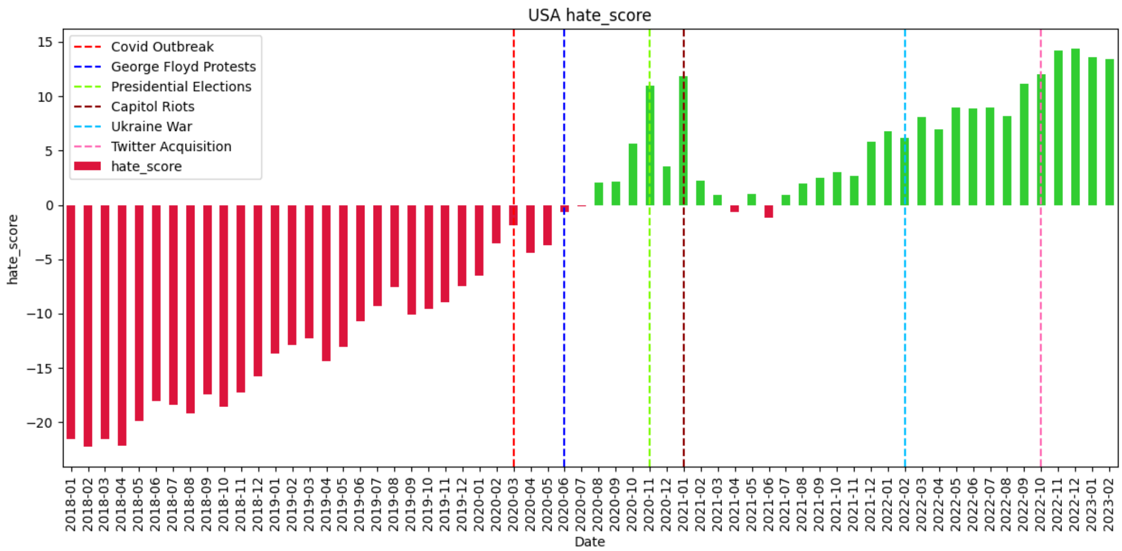
Additionally, there's a pronounced surge in late 2022, aligning with Elon Musk's acquisition of Twitter—a move that was controversial among users.



Regarding, imposed COVID-19 restrictions, we can see that the graph presents a correlation between extended stay-at-home policy and increase in offensive tweets.

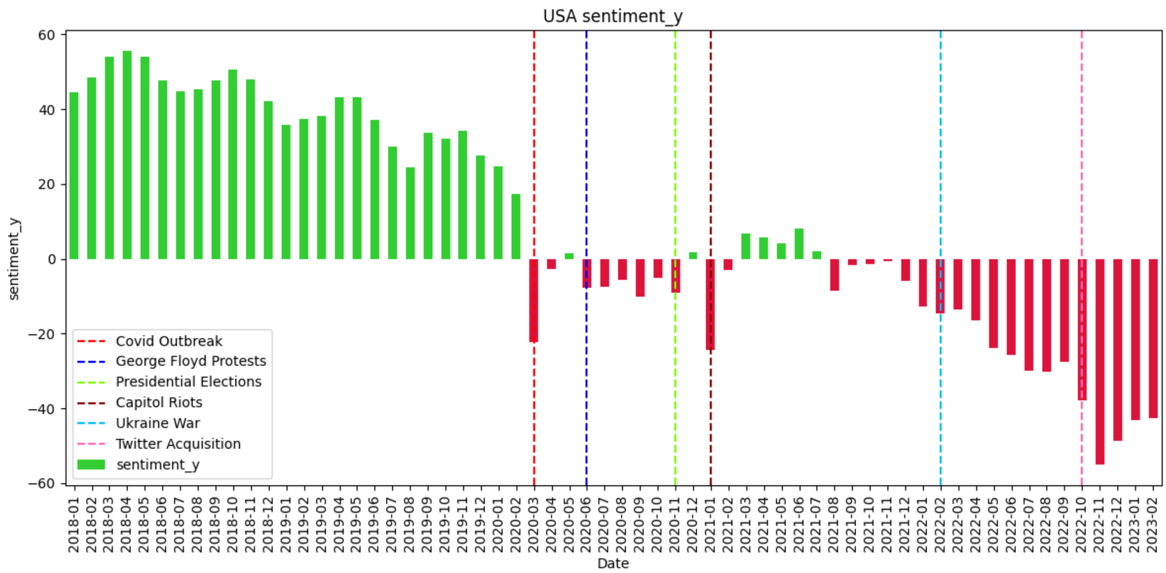
The correlation between extended stay-at-home policies and the increase in offensive tweets can be observed during the pandemic. This rise may be attributed to the stress and frustration caused by prolonged indoor confinement, disruptions to daily life, and the overall uncertainty. According to Biliotti et al. (2022), social isolation and fewer face-to-face interactions during lockdowns could lead to fewer inhibitions in online communication, resulting in more offensive or confrontational language, particularly on platforms like Twitter.

HATE



The hate and offensive indices seem to perform very similarly, with the graph showing an overall increase in hate score on Twitter. Peaks in hate speech correspond with the resurgence of COVID-19 cases and associated restrictions in late 2020, and are further pronounced during major political events in the U.S. during November 2020 and January 2021. These spikes suggest a correlation between societal stressors and heightened expressions of hate online.

SENTIMENT



The sentiment index depicted in the graph captures a significant shift in the collective mood on Twitter over the studied time period. Although a downward trend in sentiment can be observed before the start of the pandemic, there is a sharp decline in positive sentiment at the onset of the COVID-19 pandemic, mirroring the global uncertainty and fear during that period.

January 2021 marks another steep decline, likely a reflection of the political climate at the time. Sentiment then briefly returns to its average values before plunging at the beginning of 2022.

The sentiment index reaches its lowest point in November 2022, suggesting that the cumulative effect of ongoing global issues such as the war in Ukraine and significant events such as Elon Musk's acquisition of Twitter are manifesting in a particularly negative public mood at this time. This all-time low in sentiment underscores the long-term emotional toll that successive crises may have on the society.

STATISTICAL ANALYSIS

Following the examination of the emotional trends depicted in the preceding graphs, this section delves into the statistical analysis of these changes. By comparing the average emotional indices before and after the COVID-19 pandemic, we employ t-tests to assess the significance of these shifts.

The date Match 1st 2020 was chosen as the starting point of COVID-19 as it closely aligns with the global spread of the coronavirus and the World Health Organization's declaration of a pandemic on March 11th 2020 .

Emotion	Before Mean	After Mean	% Difference	t-Statistic	p-Value
Fear	0.040	0.041	2.50%	-14.353	0.000
Anger	0.161	0.202	25.47%	-485.489	0.000
Joy	0.410	0.371	-9.51%	317.455	0.000
Optimism	0.282	0.256	-9.22%	281.424	0.000
Sadness	0.148	0.171	15.54%	-314.530	0.000
Offensive Score	0.106	0.123	16.04%	-399.379	0.000
Hate Score	0.038	0.047	23.68%	-545.346	0.000
Sentiment	0.347	0.205	-40.92%	499.300	0.000

Expanding on the statistical verification of emotional changes using t-tests, our results show notable shifts in emotions before and after the start of the COVID-19 pandemic.

Among all the emotions, sentiment experienced the most significant change, with a steep decline of 40.92%. This substantial decrease in overall sentiment highlights the psychological pressure and a generally more pessimistic outlook during and after the pandemic. Furthermore, our analysis indicates a notable increase in negative emotions. Anger (25.47%), Hate (23.68%), Sadness (15.54%), and Offensiveness (16.04%) all showed significant upward trends. Other emotions such as Joy (down by 9.51%), optimism (down by 9.22%), and fear (up by 2.5%) showed smaller but still significant shifts.

Each of these changes is statistically significant, indicating a substantial change in the emotional environment brought about by the pandemic. The rise in negative emotions and the significant decrease in overall positive sentiment vividly illustrate the profound impact of the pandemic on mental health and well-being.

However, it's essential to acknowledge that examining the entire range of dates for this analysis might not provide the most accurate representation, as other significant events occurred during the COVID-19 pandemic that had an impact on the populations well-being. Therefore, we focused specifically on the period immediately before and after the onset of the pandemic to capture the most direct impact of COVID-19 on emotional trends.

To capture the immediate emotional shift triggered by the COVID-19 pandemic, a comparative analysis was conducted on the average emotion index from the immediate pre-pandemic period (December 1st, 2019, to February 29th, 2020) with that from the initial phase of the pandemic (March 1st to May 31st, 2020).

By isolating this six-month period, the comparison is designed to quantify the immediate impact of the pandemic on public sentiment, without interference from other major events of the time, such as the Black Lives Matter protests in June 2020.

To statistically validate the differences observed in the emotional indices, an independent samples t-test was used. This method provided a rigorous assessment of whether the changes in emotions such as fear, joy, sadness, and anger were significant.

Emotion	Before Mean	After Mean	% Difference	t-Statistic	p-Value
Fear	0.039	0.049	24.76%	-77.836	0
Anger	0.173	0.182	5.22%	-32.67	0
Joy	0.4	0.36	-10.00%	102.575	0
Optimism	0.271	0.27	-0.35%	3.21	0.001
Sadness	0.156	0.188	20.45%	-126.05	0
Offensive Score	0.11	0.115	4.18%	-33.963	0
Hate Score	0.042	0.043	2.53%	-19.56	0
Sentiment	0.304	0.227	-25.32%	83.618	0

The data analysis reveals the significant emotional impact of the pandemic, with increases in all negative emotions and decreases in positive ones as a response to COVID-19. The most profound shifts are a notable increase in fear (up by 25%) and sadness (up by 20%), highlighting the distress and sorrow caused by the uncertainties and losses of the pandemic. Other emotions like anger, hate, and offensive content, while not as dramatically affected as fear or sadness, also register increases.

The overall sentiment, representing the collective mood, has plummeted by 25%, marking the most substantial shift among the measured emotions and encapsulating the general downturn in positivity. Similarly, joy also decreased by 10% reflecting a reduced ability to experience happiness and pleasure during the early stages of the global crisis.

Strikingly, optimism appears almost unchanged, with a decline of just 0.35%. This relative steadiness might seem unexpected amid the negative atmosphere. It is important to note, however, that this slight but statistically significant decrease (p-value = 0.001) continues a pre-existing downward trend in optimism observed before the pandemic. Despite its significance, the magnitude of change in optimism is considerably less dramatic compared to other emotions, suggesting that it was less impacted by COVID-19.

Overall, the data depicts a clear emotional shift towards negativity and stress in response to the pandemic, with positive emotions and sentiment taking the hardest hit, while traditionally resilient feelings like optimism exhibit only minor changes.

The emotional responses observed over the different periods of analysis underscore the complex nature of how global crises, such as the COVID-19 pandemic, influence public sentiment. Specifically, the initial surge in fear and sadness can be attributed directly to the immediate shock and uncertainty triggered by the pandemic's onset. This sharp increase reflects the instinctive human response to sudden, disruptive changes and the direct threats to health and stability posed by the pandemic.

In contrast, the emotional metrics of hate score, offensive score, and optimism displayed bigger differences when observed over the extended timeframe. This suggests that these emotions might be less reactive to sudden disruptions but are more susceptible to a cumulative effect of sustained stress and societal shifts. The slower build-up in these emotions indicates a different kind of resilience, where the initial impact of the pandemic might not have caused a large spike, but prolonged conditions of uncertainty contributed to their gradual rise.

Our analysis shows that over a longer period, emotions like anger and hate seem to be influenced more by ongoing political and social events than by COVID-19 itself. These emotions did not just spike suddenly due to the pandemic; rather, they gradually increased as the political and social climate continued to evolve, which was influenced by, but not limited to, the pandemic. Understanding how emotions like these change immediately and over time helps us see which aspects of public sentiment are more resilient or vulnerable to different kinds of disruptions.

FIXED EFFECTS REGRESSIONS

In this segment of our analysis, we have focused on utilizing fixed effects regressions to better understand how emotions have changed because of the crisis. We are using fixed effects regressions to particularly examine how each emotion varied across individual users over time.

To achieve a robust understanding of these dynamics, we have selected each individual emotional metric as a dependent variable, while tweet characteristics such as retweets, mentions, hashtags, urls, and favorites, which can significantly influence a tweet's visibility and the way it is perceived, are used a control variables.

Additionally, we incorporate an instrumental variable (IV) to identify the impact of the COVID-19 period. This variable is coded as 1 for tweets posted during the COVID-19 pandemic and 0 before it, allowing us to isolate the effect of the pandemic from other temporal effects.

The fixed effects regressions completed during this part, take into account the whole span of our data (Jan 2018 – Feb 2023). Similarly to the previous analysis, the starting point of COVID-19 is considered Mar 1st 2020.

	coefficient	t_value	p_value	significance
Fear	0.004	20.511	0.000	***
Anger	0.021	42.933	0.000	***
Joy	-0.034	-52.984	0.000	***
Optimism	0.001	2.285	0.022	*
Sadness	0.012	28.103	0.000	***
Sentiment	-0.060	-38.706	0.000	***

Offensive score	0.006	24.518	0.000	***
Hate score	0.000	5.609	0.000	***
Signif. codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1**				

INTERPRETATION OF RESULTS

Fear: A coefficient of 0.004 with a t-value of 20.511 and a highly significant p-value indicates that fear increased significantly during the pandemic.

Anger: The positive coefficient of 0.021 and a t-value of 42.933, both highly significant, reflect a considerable increase in anger.

Joy: A negative coefficient of -0.034 and a t-value of -52.984, both significant, imply a substantial decrease in joy.

Optimism: Although the coefficient is small (0.001), its significance (p-value: 0.022) suggests a measurable but modest change in optimism.

Sadness: With a coefficient of 0.012 and a t-value of 28.103, sadness shows a significant increase, aligning with the general downturn in emotional well-being.

Offensive Score: A coefficient of 0.006, along with a high t-value, signifies a noticeable increase in offensive content.

Hate Score: A coefficient near zero (0.000) but still significant with a t-value of 5.609 suggests a minor increase in expressions of hate, which might reflect heightened societal divisions during the pandemic.

Sentiment: The negative coefficient of -0.060 and a t-value of -38.706 indicate a significant decrease in overall positive sentiment, reflecting the severe impact of the pandemic on general mental health.

The fixed effects model shows clear evidence that the pandemic significantly influenced individuals. Notably, negative emotions such as fear, anger, and sadness significantly increased, while positive emotions and general sentiment declined sharply.

By focusing on individual users, we have gained insights into how specific individuals' emotional expressions have shifted during the COVID-19 pandemic. This approach contrasts with previous analyses that aggregated emotional trends across the whole population, having the benefit of providing a more personalized view.

We observed a significant increase in both hate scores and offensiveness, indicating that individuals are expressing more hostility in their tweets. Simultaneously, the noticeable decline in joy suggests that there is a reduction in positive content shared by users. These results show a critical shift in the language and content shared on Twitter, where the individual emotional climate has become markedly more negative.

This individual-level analysis emphasizes the profound impact of the pandemic on personal expression and interaction on Twitter, with users opting for more expressions of hostility, such as offensive content and hate speech. This trend highlights a significant shift in the nature of interactions on social media following the COVID-19 pandemic and measures that accompanied it.

This shift in individual expressions is not only a reflection of the pandemic's psychological toll but also aligns with the broader emotional trends identified in our earlier aggregate analysis. The correlation between these findings underscores the importance of monitoring and addressing the growing prevalence of negative interactions on social platforms.

4.3 CASE STUDY: SWEDEN

APPROACH TO THE COVID-19 PANDEMIC

In contrast to many nations' responses to the COVID-19 pandemic, Sweden followed a strategy of minimal compulsory measures and a strong reliance on individual responsibility and voluntary public health measures. This approach, marked by minimal mandatory measures and a reliance on trust between the government and its citizens, sparked significant discussion and debate both at home and abroad. In this section, we aim to analyze Sweden's approach to the COVID-19 pandemic, exploring its defining characteristics, controversies, and outcomes.

The following detailed analysis of Sweden's approach to the COVID-19 pandemic is based on insights from R. Chen's paper, "A timeline of events," published in December 2022, which provides an overview of the critical events during the pandemic in Sweden (Chen, 2022).

No Formal Lockdowns

Sweden chose not to implement formal lockdowns. Instead, the strategy was to rely on the public's sense of responsibility and adherence to voluntary measures, and in that way avoid the economic and social disruptions associated with strict lockdowns.

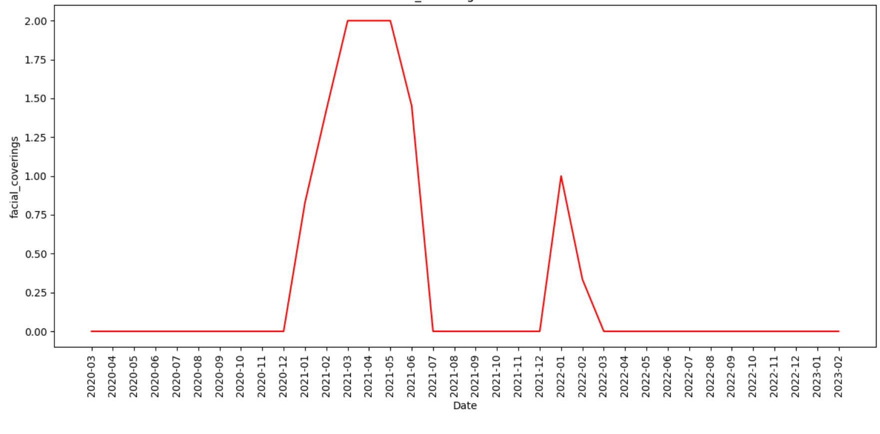
Schools and Educational Institutions

One of the most significant decisions was to keep educational institutions open for younger children throughout the pandemic. This decision was underpinned by the belief that closing schools would not significantly decrease the virus spread but would negatively impact children's education and well-being.

Recommendations Versus Regulations

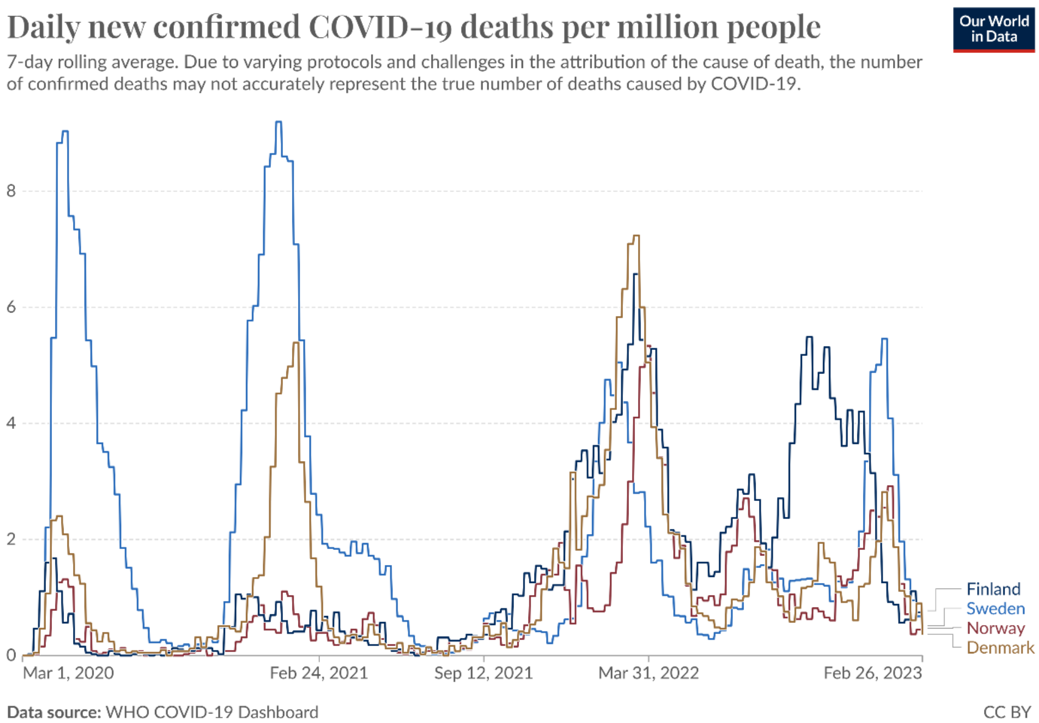
Sweden's approach emphasized recommendations over legal regulations. The Public Health Agency of Sweden (FHM) promoted remote working, social distancing, and avoiding non-essential travel, rather than making them legal requirements.

Limited Use of Face Masks



Initially, Sweden did not recommend the use of face masks in public. This stance was maintained well into the pandemic, based on assessments that the evidence regarding mask effectiveness was inconclusive. It was not until the end of 2020 that the guidelines were adjusted to recommend masks in crowded settings, such as public transport during peak times.

Criticism and Defense of Sweden's Approach



The passive strategy received a significant amount of criticism, particularly regarding its effectiveness as the pandemic unfolded. Critics argued that Sweden's higher mortality rate, compared to its Nordic neighbors, illustrated the shortcomings of its pandemic response. However, supporters, including the state epidemiologist Anders Tegnell, defended the strategy by suggesting that stringent lockdowns were not sustainable over the long term and could lead to more severe secondary impacts on society and the economy.

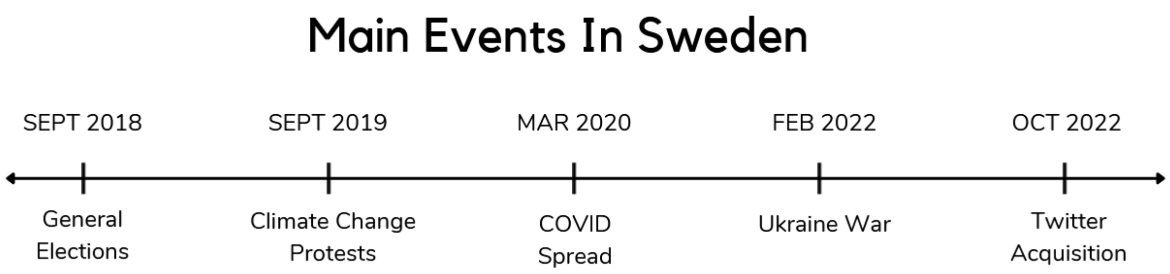
Long-Term Perspective

Swedish officials often described their pandemic response as a long-term strategy, likening it to a marathon rather than a sprint. This metaphor underscored their belief that the measures implemented could be sustained over time without the drastic societal and economic costs associated with more aggressive strategies.

Due to the significantly different policy adopted by Sweden, especially in contrast to the more stringent measures taken by other countries, like the USA, it has been chosen for further analysis. This analysis will first examine Sweden's sentiment fluctuations on Twitter and then compare them to those of the USA, providing a comparative perspective on the impacts of varied public health strategies during the pandemic.

MAIN EVENTS

In examining public sentiment, it is crucial to consider major events in Sweden that could have influenced it, even in a nation known for its relative peace and less stringent policy approaches. While Sweden often adopts more measured stances in its domestic and foreign policies, significant events, such as the Ukraine War, have the potential to deeply impact public sentiment and discourse. In this section we will look at the most prevalent events in Sweden for the studied time period.



September 2018 - General Elections

In September 2018, Sweden held its general elections, which marked a significant political moment for the country. The elections saw the rise of the far-right Sweden Democrats party, which gained ground. The outcome of these elections, with no single party gaining a clear majority, led to a period of political uncertainty and negotiations to form a coalition government (Aylott & Bolin, 2019).

September 2019 - Climate Change Protests

In September 2019, Sweden experienced widespread climate change protests, inspired by the global youth movement led by Greta Thunberg, a Swedish environmental activist. These protests were part of a worldwide movement demanding urgent action on climate change, coinciding with the 2019 UN Climate Summit. In Sweden, thousands of people, particularly the youth, took to the streets in various cities, signaling a strong public response to environmental issues. The scale and passion of these protests would have had a noticeable impact on social media sentiments, reflecting a heightened awareness and concern for climate issues. (Boulianne, S. et al, 2020)

February 2022 - War in Ukraine

Similar to its effects in the USA, the Ukraine War also significantly impacted Sweden, although the country's response was tailored to its unique position of neutrality and its proximity to Russia. The conflict raised concerns about security in Sweden, and sparked political discussions about the possibility of joining NATO. This shift in security policy perspective became a prominent subject on social media, reflecting both concerns for Ukraine and Swedish national security. (TODAY, 2024)

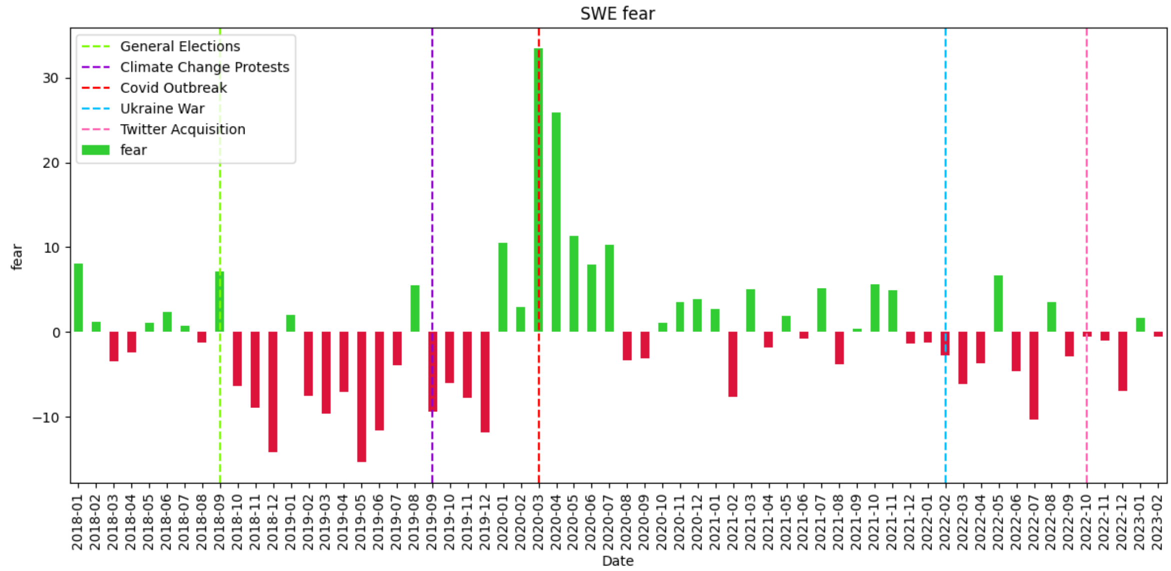
October 2022 - Acquisition of Twitter

The acquisition of Twitter by Elon Musk resonated in Sweden much as it did in the United States. In Sweden, a country known for high social media use, the uncertainty about Twitter's future governance sparked discussions about digital rights, privacy, and the platform's role in Swedish public and political communication.

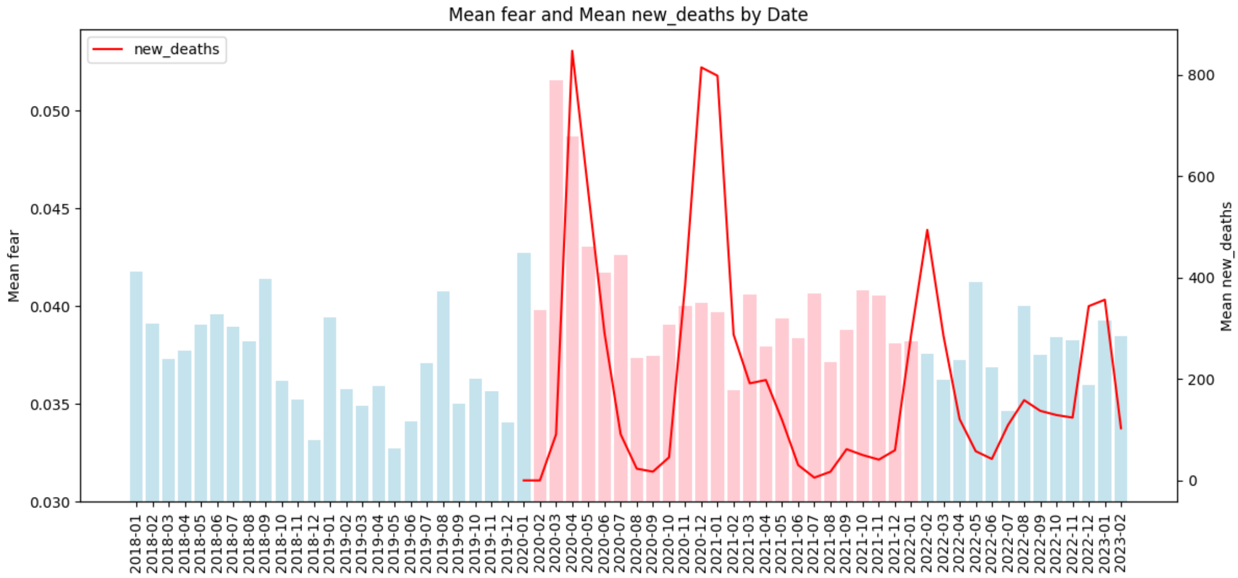
SENTIMENT FLUCTUATIONS

The following analysis provides a comprehensive view of the emotional landscape of Sweden, as interpreted through Twitter data. By examining the percentage change in average monthly emotions alongside the mean emotional index, we can understand how public sentiment has fluctuated over time. Through this lens, we can better comprehend the emotional reactions of Swedish Twitter users to global crises, policy changes, and other significant occurrences during the studied period.

FEAR

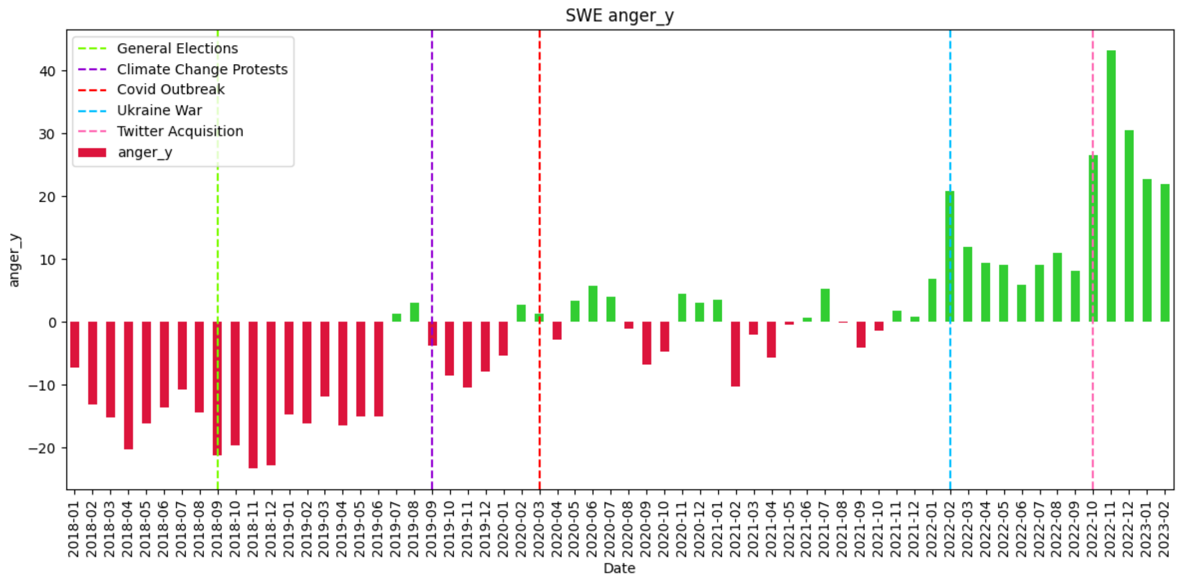


The analysis of fear in Swedish tweets reveals significant fluctuations tied to major events, with the sharpest increase being noted during the initial phase of the COVID-19 pandemic. Starting in January 2020, fear surged likely due to early reports of the virus in China and remained high until June 2020 when the cases started to decrease. COVID-19 related spikes can be seen again in November 2020, aligning with the second wave of COVID-19, reflecting widespread anxiety about health and economic stability. Additionally, other events like the general elections in September 2018 also saw spikes in fear, potentially due to political uncertainties. International incidents, such as the Ukraine conflict in February 2022 and the acquisition of Twitter in October 2022, while influential, appear to not have caused fear amongst Swedish Twitter users.



The above graph displays the surge in fear that can be tied to the second wave of the pandemic.

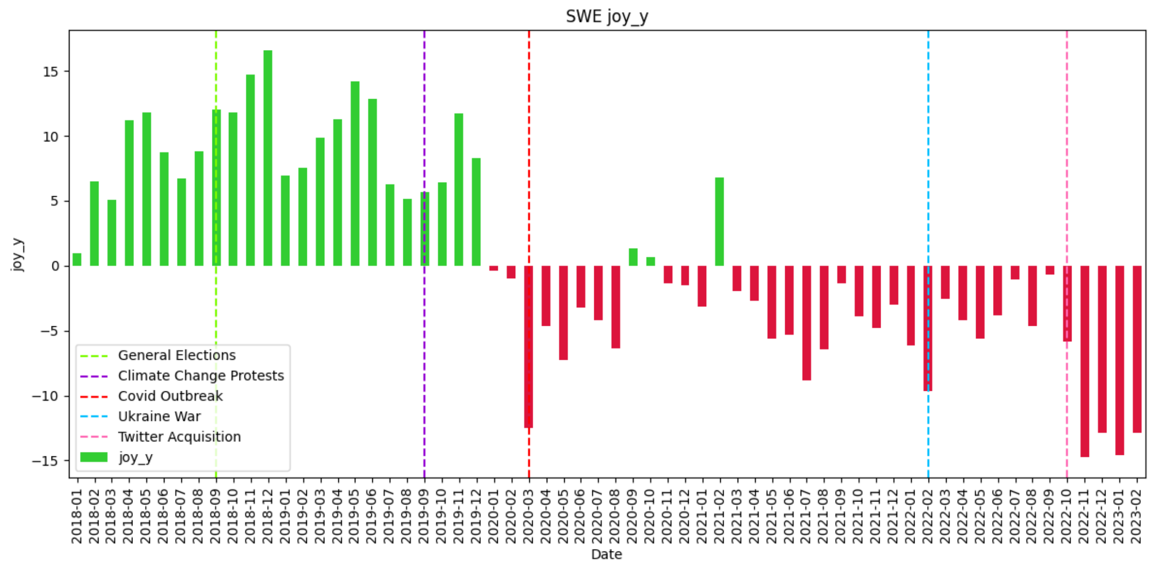
ANGER



The analysis of the monthly percentage change in anger in Swedish tweets indicates that despite the global upheaval caused by COVID-19, anger levels in Sweden show only a slight uptick during this period, suggesting that the pandemic alone did not significantly elevate anger on social media within the country. However, there is a pronounced increase in expressions of anger in 2022, particularly noticeable with two significant spikes. The first surge in February 2022 coincides with the onset of the Ukraine War, reflecting public frustration. A second spike in October 2022 aligns with the news of Elon Musk's acquisition of Twitter, which provoked strong reactions worldwide. These increases suggest that while the pandemic's direct impact on anger was muted, other global and significant events triggered more substantial expressions of anger among the Swedish Twitter users.

It's noteworthy that the relatively stable levels of anger during the early stages of the COVID-19 pandemic in Sweden could be attributed to the country's unique approach to pandemic management. Unlike many other countries, Sweden did not impose lockdowns or harsh restrictions until later in the pandemic, promoting a sense of freedom among its citizens. This strategy might have contributed to a perception of autonomy and normalcy, potentially mitigating feelings of anger or frustration that were more prevalent in countries with strict lockdown measures (as seen in the USA for example). This context of perceived freedom could explain why the Swedish population displayed less anger on social media during the initial months of the global health crisis.

JOY

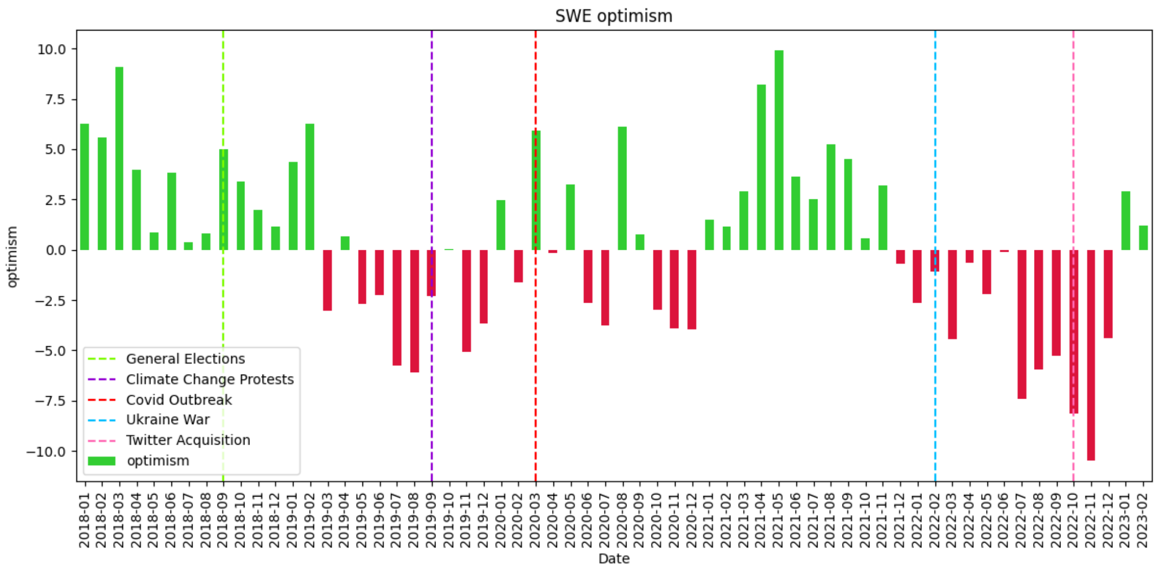


Starting in early 2020, there is a decline in joy that correlates with the initial international spread of the coronavirus, particularly as the situation worsened in China. This early decrease suggests that Swedish Twitter users were quick to react emotionally to the global uncertainty and potential threat posed by the new virus, even before

COVID-19 had a significant direct impact on Sweden. Joy is observed to drop further in the following months as COVID-19 starts spreading in Sweden.

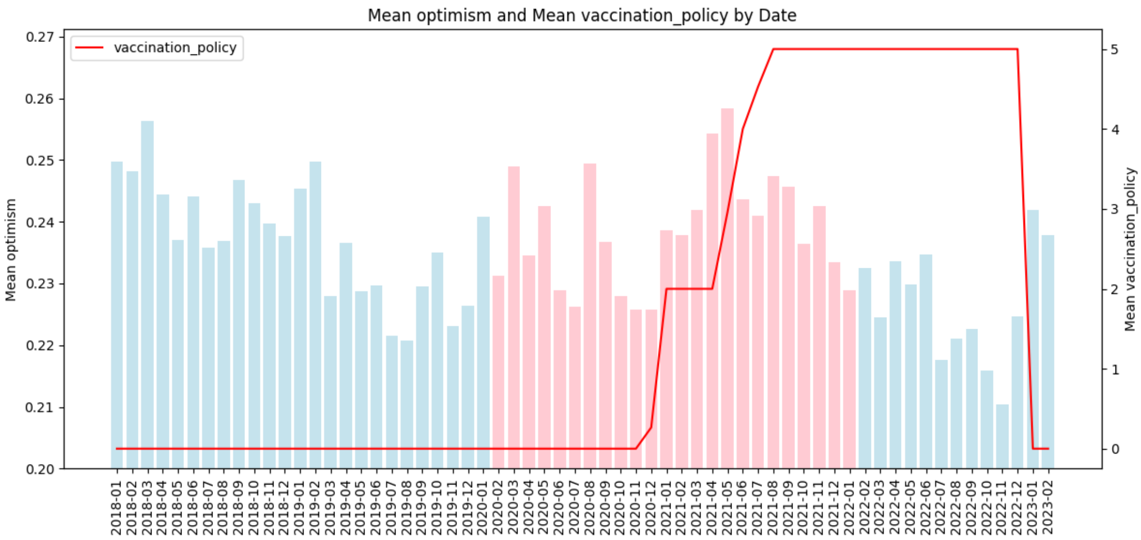
Following this initial drop, joy levels on Twitter, although increasing a little bit after the first wave, do not return to their prior levels, indicating a sustained impact of the pandemic on positive expressions Additionally, a notable decrease in joy is observed around Russia’s invasion of Ukraine and Elon Musk's acquisition of Twitter.

OPTIMISM



The graph displaying the monthly percentage change in optimism in Swedish tweets offers an interesting perspective. Notably, optimism does not seem to be negatively affected by the initial outbreak of COVID-19, with optimism levels in March 2020 being the highest recorded in the previous 12 months.

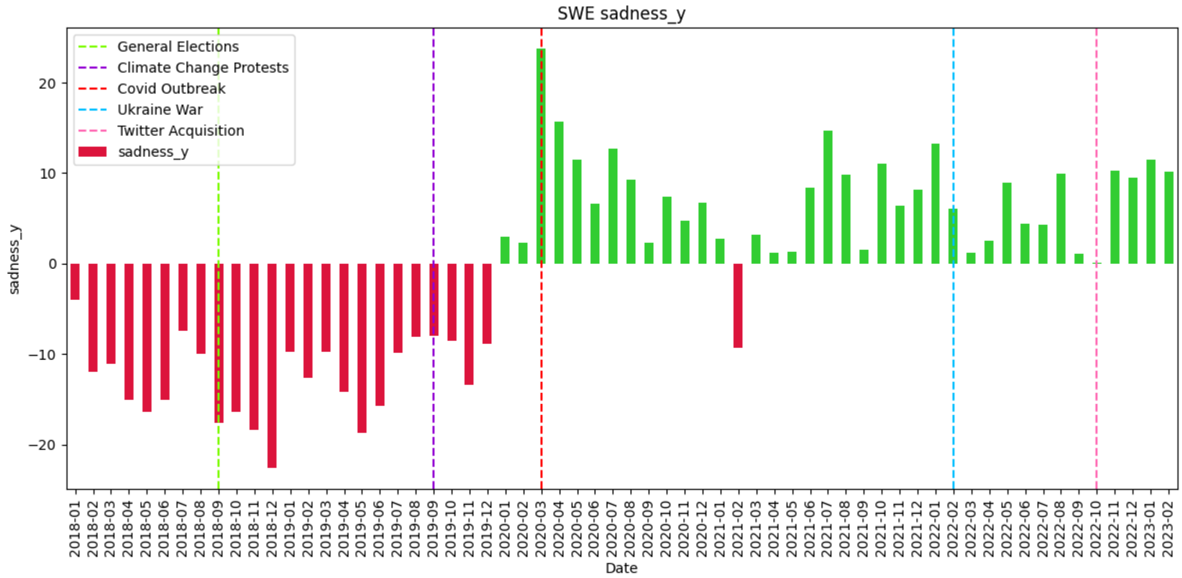
Other events however, such as the climate protests in 2019, Russia's invasion of Ukraine in February 2022, and the Twitter acquisition in October 2022 all coincided with lower-than-average optimism scores. These observations suggest that Swedish Twitter users' levels of optimism are particularly sensitive to international events that introduce uncertainty and potential threats, rather than domestic policy actions or health crises alone.



The graph depicting the mean optimism and vaccination policy in Sweden reveals that at with the start of Sweden's vaccination campaign, there is a sharp increase in optimism. This increase likely reflects a collective sense of relief and hope among Swedish Twitter users and hope for a potential return to normalcy.

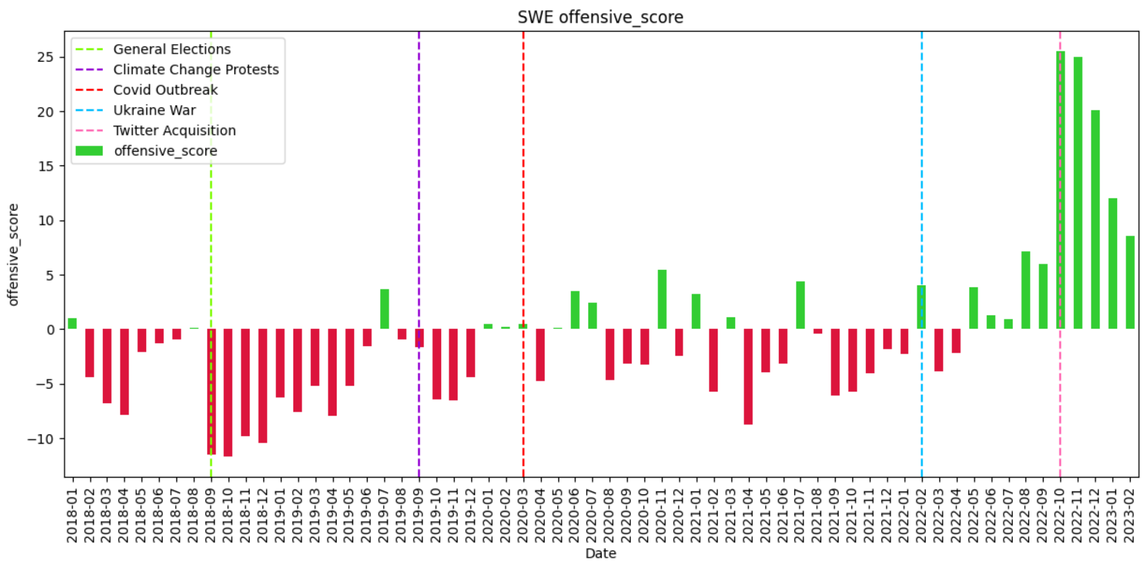
No other significant events were recorded during this period that could account for this spike in positive tweets, suggesting that the vaccine rollout itself was a primary driver of the improved sentiment. This observation shows the impact of health-related developments on public emotion and highlights how pivotal moments in a pandemic, such as the start of vaccination, can lift public spirits and drive positive social media expressions.

SADNESS



The analysis of sadness in Swedish tweets reveals a notable increase beginning early 2020, peaking in March as the COVID-19 pandemic escalated globally. Sadness levels then gradually decline towards the end of 2020, however, sadness rises again and stabilizes at high levels through 2021, reflecting ongoing pandemic challenges. A brief but significant decrease in sadness occurs in February 2021, coinciding with the start of Sweden's vaccination campaign.

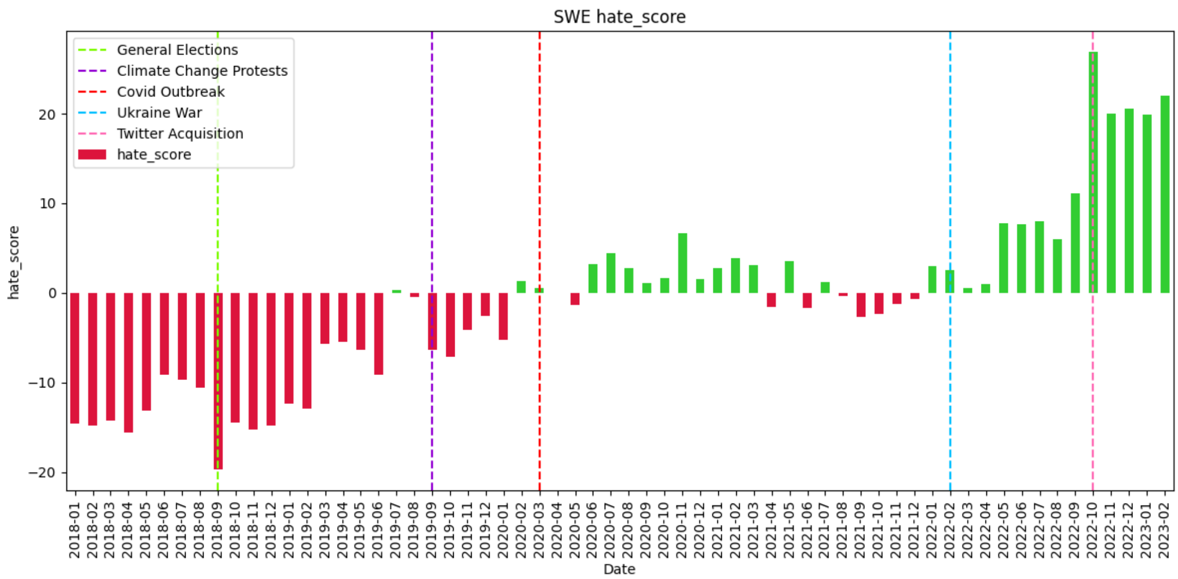
OFFENSIVENESS



The graph tracking the monthly percentage change in offensiveness in Swedish tweets shows a relatively steady level through the duration of the COVID-19 pandemic, suggesting that the health crisis did not significantly affect the frequency of offensive postings. However, a sharp increase in offensiveness is observed towards the end of 2022, which aligns with the acquisition of Twitter by Elon Musk.

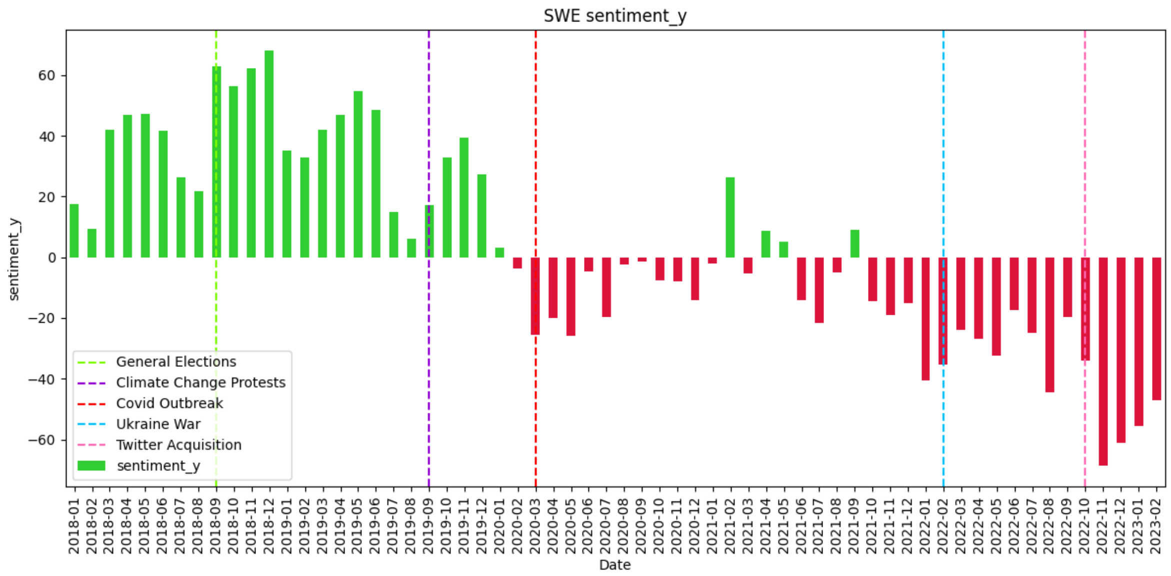
The stability in offensive content in Swedish tweets during the early stages of the COVID-19 pandemic can be explained similarly to the trends observed in anger expressions. Sweden's approach to the pandemic, characterized by fewer restrictions and more reliance on voluntary compliance rather than strict lockdowns, likely contributed to maintaining a lower level of offensive tweets.

HATE



The analysis of hate speech in Swedish tweets shows a distinct upward trend, which can be separated into three periods: Before COVID-19, where hate levels were consistently below average, indicating stable minimal expressions of hate; after the onset of COVID-19, where there is a slight increase, likely due to heightened anxieties and social tensions from the pandemic's impacts on health and economy; and lastly around the Twitter acquisition, where hate speech significantly spikes.

SENTIMENT



This graph, depicting the monthly percentage change in overall sentiment, shows a noticeable downward trend. Initially, there is a decline in sentiment at the beginning of the COVID-19 pandemic, reflecting the global uncertainty and fear associated with the virus. This trend is consistent with the observed increases in negative emotions such as fear and sadness during the same period.

A further, gradual decrease in sentiment is observed starting in early 2022, with November 2022, recording the lowest levels recorded. This drop coincides with significant global and local events, including the ongoing Ukraine War and the acquisition of Twitter by Elon Musk, both of which also correspond to spikes in anger and hate speech as seen in previous analyses. The culmination of negative reactions to these events likely contributed to a very pessimistic overall sentiment.

These findings of a declining sentiment align with the patterns seen in the emotional analyses described above — where joy decreases, while fear, hate, and sadness increase, particularly during key events of global significance. These emotional trends offer insights into the general climate of Swedish Twitter users over the studied years, highlighting the influence of global challenges and on public sentiment.

Despite the observation that specific emotions such as anger, offensiveness, and optimism appeared less impacted by COVID-19 due Sweden's less restrictive pandemic policies, the overall sentiment among Swedish Twitter users has shown a consistent downward trend. This decline suggests that while the government's approach may ease the emotional burden of the pandemic, the cumulative effect of the pandemic still led to a general decline in public sentiment.

STATISTICAL ANALYSIS

After reviewing the emotional patterns fluctuations in the graphs above, this section proceeds to analyze these changes statistically. We use t-tests to evaluate the significance of the differences in average emotional indices before and after the onset of the COVID-19 pandemic.

March 1st 2020 was selected as the starting date for COVID-19, as with all other analyses from this study.

Emotion	Before Mean	After Mean	% Difference	t-Statistic	p-Value
Fear	0.037	0.039	5.41%	-5.664	0.000
Anger	0.147	0.180	22.45%	-38.888	0.000
Joy	0.473	0.413	-12.68%	45.531	0.000
Optimism	0.237	0.234	-1.27%	4.095	0.000
Sadness	0.143	0.174	21.68%	-37.836	0.000
Offensive Score	0.100	0.108	8.00%	-17.814	0.000
Hate score	0.040	0.046	15.00%	-40.314	0.000
Sentiment	0.339	0.196	-42.18%	46.970	0.000

The statistical analysis reveals varying degrees of impact across different emotions, with the most pronounced changes observed in sentiment, anger, and sadness. Sentiment experienced the most significant decline, dropping by 41.18 points, reflecting a substantial negative shift. This in combination with increases in anger and sadness by 22.45% and 21.68% respectively, indicates heightened anxiety and frustration among the population.

Similarly, hate and joy were also significantly affected, with hate increasing by 15% and joy decreasing by 12.68%. These changes suggest a noticeable shift towards more negative emotional expressions, with reduced experiences of positive emotions like joy during the pandemic period.

Emotions such as offensiveness and fear were moderately affected (8% and 5.41% respectively), while optimism showed the least variation, with only minor but still significant changes. The analysis points to a general trend where negative emotions like anger, sadness, and hate tend to increase, while positive emotions such as joy and overall sentiment show declines over the full span of the data from 2018 to 2023. This trend underscores a shift towards a more negative emotional landscape on Swedish Twitter through the duration of the pandemic.

To effectively understand the immediate emotional impact of the COVID-19 pandemic on Swedish Twitter users, a focused comparative analysis over a six-month period was conducted. This analysis specifically measured the average emotion index from the immediate pre-pandemic period (December 1st, 2019 - February 29th, 2020) against the average emotion index from the initial phase of the pandemic (March 1st - May 31st, 2020). By narrowing the timeframe to these critical six months, the analysis aims to capture the immediate changes in public sentiment as the pandemic began to unfold.

This approach also allows for making direct parallels to the analysis done for the USA, allowing for a direct comparison, and offering insights into how different national responses may have influenced public sentiment differently.

Emotion	Before Mean	After Mean	% Difference	t-Statistic	p-Value
Fear	0.039	0.048	23.08%	-6.864	0
Anger	0.161	0.168	4.35%	-2.419	0.016
Joy	0.445	0.4	-10.11%	10.857	0
Optimism	0.233	0.242	3.86%	-3.153	0.002
Sadness	0.16	0.19	18.75%	-10.995	0
Offensive Score	0.104	0.103	-0.96%	0.156	0.876
Hate Score	0.043	0.044	2.33%	-1.53	0.126
Sentiment	0.272	0.191	-29.78%	8.397	0

In comparison to the analysis performed over the whole timeframe, this analysis focused on the pandemic's early months and highlighted a different kind of shifts in emotions. For instance, the decrease in overall sentiment during these few months (-29.78%) was less pronounced than the longer-term trend (-42.18%), underscoring the long-term impact of the pandemic. Similarly, anger saw a long-term rise of 22.45% compared to a short-term increase of just 4.35% (with a relatively high p-value of 0.016), and sadness increased by 21.68% in the long-term compared to 18.75% in the short-term. Other emotions, such as fear increased significantly more in the first months of the pandemic (23.08% short-term in comparison to 5.41% long term change), highlighting an initial surge in fear which later subsided. Conversely, joy decreased significantly in both periods, with the long-term decline in joy being slightly larger (12.68%) compared to the short-term (10.11%).

In the previous analysis from 2018-2023, both offensive and hate score indices showed slight but statistically significant increases, however, during the early phase of the pandemic, these emotions did not exhibit significant changes. This is likely due to the extended analysis capturing a period where other significant events, such as tensions surrounding the war in Ukraine and the societal impacts of the Twitter acquisition by Elon Musk, contributed to a gradual increase in negative sentiments like offensiveness and hate. These events, occurring after the initial phase of the pandemic, significantly shaped the emotional landscape, leading to noticeable increases in hate and offensiveness.

Comparison to USA results

When comparing the short-term emotional responses to the initial phase of the COVID-19 pandemic between Sweden and the USA, both countries exhibit significant emotional shifts, but there are notable differences in the magnitude and direction of these changes.

Fear: Both countries experienced a similar increase in fear, with Sweden showing a 23.08% increase and the USA a slightly higher 24.76% increase. This suggests a universally high level of anxiety and concern at the onset of the pandemic despite the difference in COVID-19 prevention policies.

Joy: The decrease in joy was almost identical in both countries, with Sweden experiencing a 10.11% decrease and the USA a 10.00% decrease. This indicates a widespread reduction in positive sentiments as the pandemic began impacting daily life.

Sadness: The increase in sadness was slightly more pronounced in the USA (20.45%) compared to Sweden (18.75%). This may reflect the more intense immediate impact of the pandemic on the US population, or differences in public health measures and media coverage between the two countries.

Anger: The USA saw a significant increase in anger (5.22%), whereas the change in Sweden (4.35%) was not statistically significant. This suggests that the pandemic may have provoked more pronounced feelings of frustration in the USA, potentially influenced by different national responses to the crisis.

Optimism: Interestingly, optimism in Sweden showed a slight increase (3.86%), whereas in the USA, it slightly decreased (-0.35%). This divergence could be attributed to the different cultural or governmental responses, fostering a sense of hope in Sweden and a slight pessimism in the USA.

Hate and Offensive Scores: Both hate and offensive behaviors saw minimal changes in Sweden, with hate increasing slightly (not statistically significant) and offensive scores actually decreasing slightly (also not significant). In contrast, the USA saw statistically significant increases in both hate and offensive scores. This could suggest that the social and political climate in the USA may have been more conducive to these expressions during the pandemic.

Overall Sentiment: The drop in overall sentiment was more severe in Sweden (-29.78%) than in the USA (-25.32%). This larger drop in Sweden could indicate a stronger reaction to the pandemic's initial phase, possibly influenced by different media narratives, public expectations, or the psychological impact of Sweden's less restrictive initial response to the pandemic.

Considerations

In analyzing the emotional responses to the COVID-19 pandemic on Twitter in Sweden and the USA, several critical factors emerge that should be taken into account:

1. Cultural Expression on Social Media

Despite the mean emotion indices for each emotion being very close between the two countries, cultural norms play an important role in how emotions are expressed on social media. The similarity in mean indices suggests that while cultural differences exist, they may not drastically change ways emotions are expressed, but they do influence the type and intensity of those expressions.

2. Reserved vs. Spontaneous Expression

The observation that hate, and offensive speech did not significantly increase in Sweden at the beginning of the pandemic could be impacted by the generally more reserved and less confrontational nature of Swedes as supported by Schramm-Nielsen's analysis on Scandinavian conflict management, which highlights an aversion to conflicts and a tendency to minimize their importance (Schramm-Nielsen, 2002). In contrast, Americans may exhibit more spontaneous and direct expressions of these emotions, which can be attributed to cultural tendencies towards more overt and immediate reactions to events that affect them.

3. Timing of Emotional Responses

In Sweden, an emotional shift attributed to COVID-19 on Twitter was noted as early as January 2020, aligning with the initial international reports about the virus. This early reaction suggests a more globally interconnected perspective, where international news have the ability to influence public sentiment. On the other hand, in the USA, increases in fear did not occur until March 2020, when the direct impacts of COVID-19 began to be felt domestically. This delay might reflect a more internally focused perspective, where international events are less likely to trigger immediate public concern until their domestic effects are observed.

The differences in emotion fluctuations surrounding the COVID-19 pandemic between Sweden and the USA can be largely attributed to the distinct approaches taken by each government in handling the crisis. However, it is crucial to recognize that these variations are not solely the result of governmental policies. Cultural norms and other societal factors also play a significant role in shaping how populations react to such global events.

FIXED EFFECTS REGRESSIONS

In this section, we used fixed effects regressions to understand how individual emotions have evolved in response to the crisis.

We're looking at each emotion separately and considering factors like favorites, retweets, mentions, hashtags, and urls as control variables. To isolate the pandemic's effect, we use an instrumental variable (IV) that marks tweets during and before the pandemic. Our analysis covers the period from January 2018 to February 2023, with March 1st, 2020 as the start of COVID-19.

	coefficient	t_value	p_value	significance
Fear	0.004	1.796	0.072	.
Anger	0.005	0.945	0.345	
Joy	-0.038	-5.036	0.000	***
Optimism	0.015	2.881	0.004	**
Sadness	0.018	3.653	0.000	***
Sentiment	-0.017	-0.943	0.346	
Offensive score	0.003	1.282	0.200	
Hate score	0.002	2.716	0.007	**
Signif. codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1**				

INTERPRETATION OF RESULTS

Fear: The coefficient is positive (0.004), indicating a slight increase in fear over time, but it's not statistically significant (p = 0.072). This suggests that while there might be a trend towards more fear, we can't be confident it's due to the pandemic.

Anger: The coefficient is positive (0.005), but not statistically significant ($p = 0.345$). This implies that there's no clear trend in anger over time among Swedish Twitter users during the crisis.

Joy: The coefficient is negative (-0.038), highly statistically significant ($p < 0.001$). This indicates a substantial decrease in joy over time.

Optimism: The coefficient is positive (0.015), statistically significant ($p = 0.004$). This suggests a modest increase in optimism during the crisis, which contrasts with the decline in joy.

Sadness: The coefficient is positive (0.018), highly statistically significant ($p < 0.001$). This points to a significant increase in sadness over time.

Offensive Score: The coefficient is positive (0.003), but not statistically significant ($p = 0.200$). This suggests no significant change in offensive language over time among Swedish Twitter users.

Hate Score: The coefficient is positive (0.002), highly statistically significant ($p = 0.007$). This indicates a notable increase in hateful language over time, highlighting a concerning trend amid the crisis.

Sentiment: The coefficient is negative (-0.017), but not statistically significant ($p = 0.346$). This suggests that there is no clear trend in overall sentiment among Swedish Twitter users during the crisis.

The analysis reveals significant emotional shifts among individual users in Sweden during the COVID-19 pandemic in joy and sadness, indicating a large impact on individual emotional experiences. Additionally, optimism and hate score also exhibited statistically significant changes, albeit to a lesser extent.

In contrast, there were no statistically significant changes in fear, anger, offensiveness, or overall sentiment for individual users in Sweden. This suggests that while internal emotional experiences such as joy, sadness, and optimism were notably affected by the pandemic, outward expressions of fear, anger, or offensive language did not significantly alter among individual users.

These findings imply that Swedish users may have experienced internal emotional shifts due to the pandemic, as shown by the changes in joy, sadness, and optimism. However, there appears to be smaller tendency among Swedish users to outwardly express these emotions towards others, as indicated by the lack of significant changes in anger, and offensiveness. This suggests a cultural tendency towards managing emotions internally rather than expressing them outwardly in the form of aggression or hostility.

Comparison to USA results

In the USA, the fixed effects model reveals a clear influence of the pandemic on individuals, with significant increases in negative emotions such as fear, anger, and sadness, alongside declines in positive emotions and general sentiment. This contrasts with Sweden, where there were no statistically significant changes in fear, anger, or sentiment, highlighting potentially different emotional responses to the crisis between the two countries. Additionally, while the USA saw significant increases in hate scores and offensiveness, Sweden did not exhibit such changes in terms of offensiveness, but did so in terms of hate.

The individual-level analysis in both countries provides valuable insights into how specific users' emotional expressions have evolved during the COVID-19 pandemic. While the USA shows a marked increase in negative interactions and a decline in positive content, Sweden's lack of significant changes in negative emotions and offensive language suggests a different pattern of emotional response.

Overall, the comparison underscores the ways in which individuals in different countries express and respond to crises like the COVID-19 pandemic. While both countries experienced emotional shifts, the specific nature and magnitude of these changes varied, reflecting cultural, social, and governmental differences in how individuals express themselves during times of crisis.

4.4 TOPIC MODELING

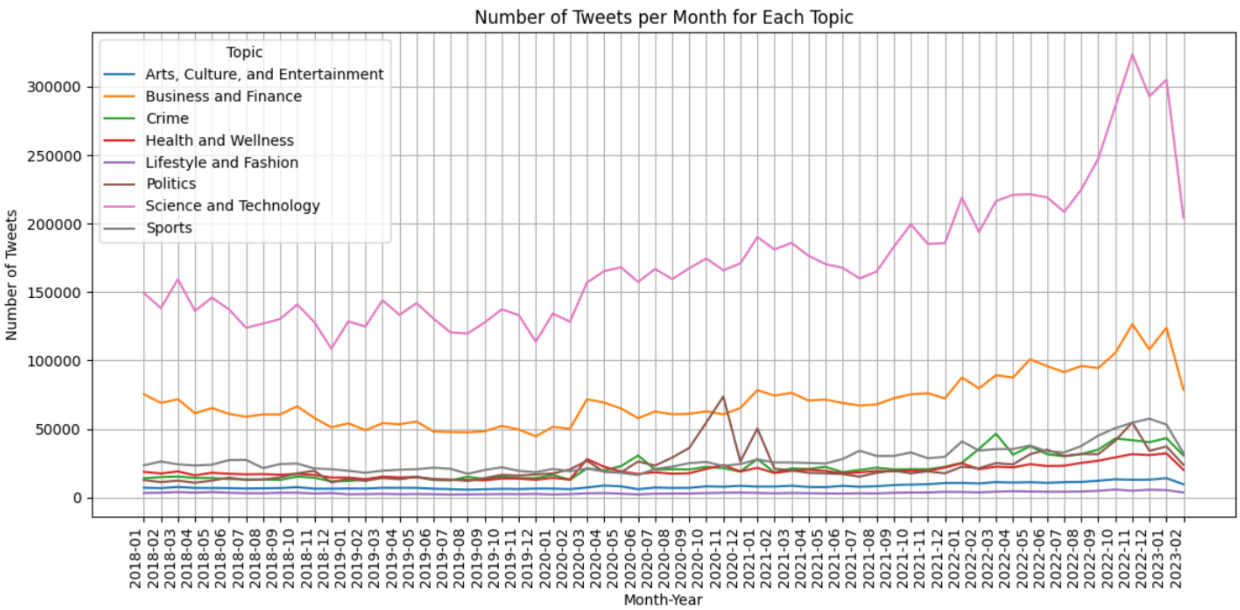
In this section of the thesis, we explore emotional fluctuations on Twitter during the COVID-19 pandemic using advanced topic modeling techniques. Employing the RoBERTa model, a variant of the BERT architecture, we have categorized tweets into eight distinct topics: Sports, Arts Culture and Entertainment, Business and Finance, Health and Wellness, Lifestyle and Fashion, Science and Technology, Politics, and Crime. We focused on tweets with topic confidence score greater than 0.9, ensuring that our analysis is based on highly relevant and accurately labeled data. Our analysis tracks the popularity and sentiment changes of these topics over time, providing insights into how public mood and opinions shifted in response to the pandemic on a global scale.

TOPIC POPULARITY

Topic	Number of tweets with topic index>0.9	Percentage of all tweets
Science and Technology	10629000	49.8%
Business and Finance	4341772	20.3%
Sports	1706526	8.0%
Politics	1399539	6.6%
Crime	1339083	6.3%
Health and Wellness	1191338	5.6%
Arts, Culture, and Entertainment	513276	2.4%
Lifestyle and Fashion	206517	1.0%

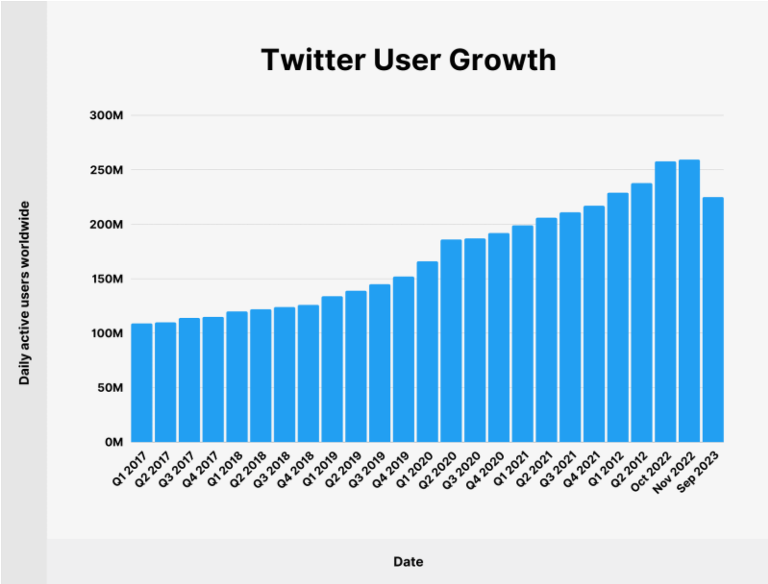
In the initial part of our analysis focusing on topic popularity, we found that Science and Technology emerged as the most discussed topic, representing 49.8% of the tweets with confidence score above 0.9 with a total of 10,629,000 tweets. This was followed by Business and Finance at 20.3%, with 4,341,772 tweets. Together, these two categories comprised 70% of the total discourse, underscoring their dominance during the period studied.

The dataset included 21,327,051 tweets in total that met our confidence criterion, with 60% of these, approximately 12,901,785 tweets, originating from the USA.



The graph provides a detailed view of how topic popularity on Twitter evolved during the studied period spanning tweets from January 2018 – February 2023, and reveals trends in public discussion.

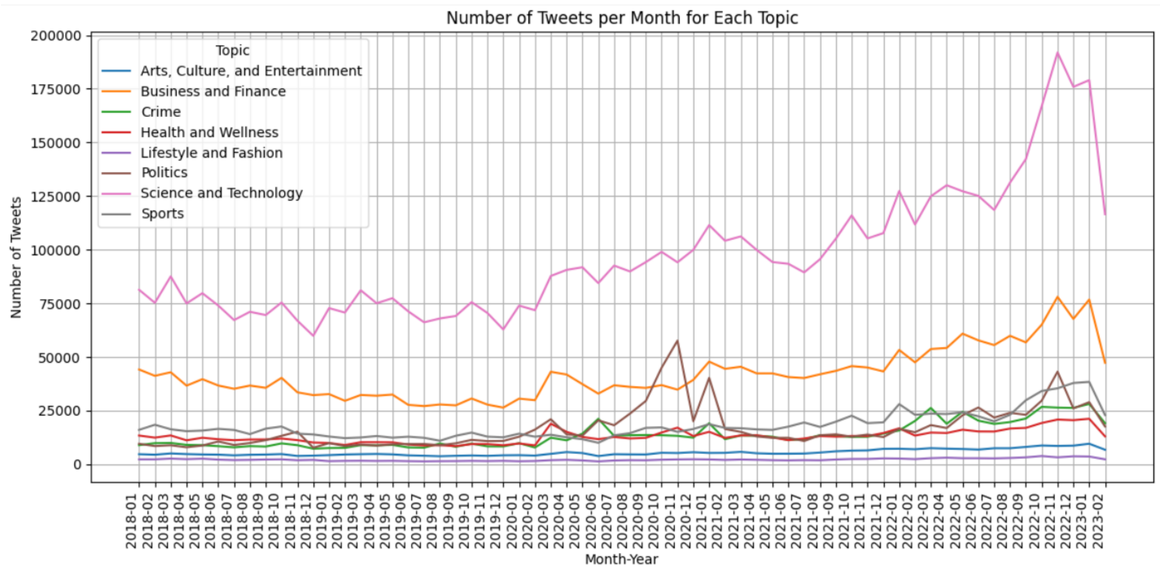
It can be observed that tweets categorized under Science and Technology have been on the rise since the beginning of the pandemic, with a sharp increase around the middle of 2022. This trend likely mirrors the increased public interest in scientific updates and technological advancements.



Source: Backlinko

Interestingly, there is a noticeable uptick in the number of tweets across all topics after mid-2022. This general increase may be attributed to the overall growth in Twitter's usage, as the platform became a key source of information and discussion worldwide.

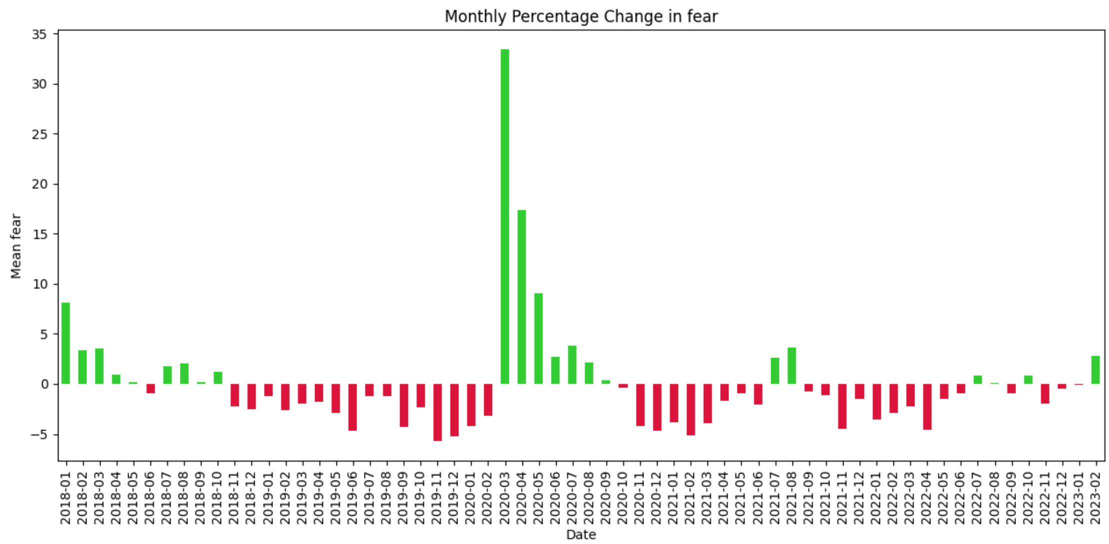
Additionally, the Politics category shows notable spikes in activity that align with major political events, such as the 2020 U.S. presidential election and the subsequent events of January 2021. Considering that 60% of all analyzed tweets originate from the USA, these spikes clearly reflect the intense engagement of American users with political content during these critical periods. The changes are more pronounced when we look at tweets originating only from the USA.



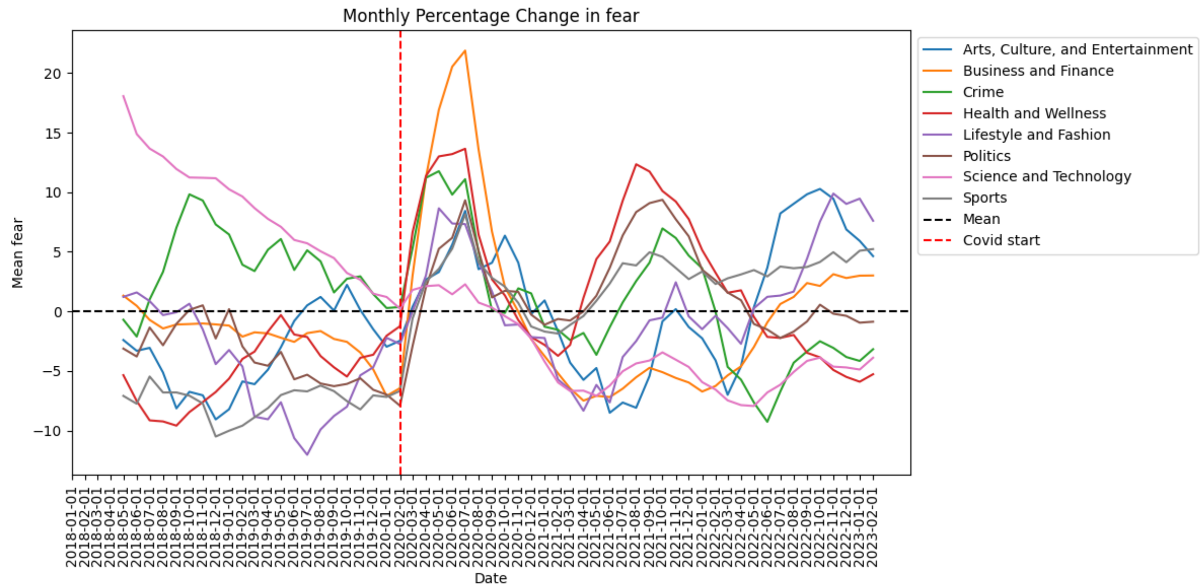
SENTIMENT FLUCTUATIONS BY TOPIC

In this section, we tracked the fluctuations of key emotions -fear, anger, joy, optimism, sadness, offensiveness, hate, and overall sentiment – across our set of aforementioned topics. The following graphs map out for each topic, the percentage difference in the mean emotion of that month in comparison to the overall mean emotion for the particular topic. This approach provides a granular view of how global crises and significant events influence public sentiment and discourse across various sectors.

FEAR



The graph showing the monthly percentage change in fear during the COVID-19 pandemic reveals a sharp initial increase as the pandemic began, indicating widespread anxiety and uncertainty. This spike in fear gradually subsides over time but shows intermittent peaks likely tied to pandemic-related developments like the spread of the Delta variant in the summer of 2021. Analyzing fear associated with specific topics will further clarify how different aspects of the crisis influenced public emotions.

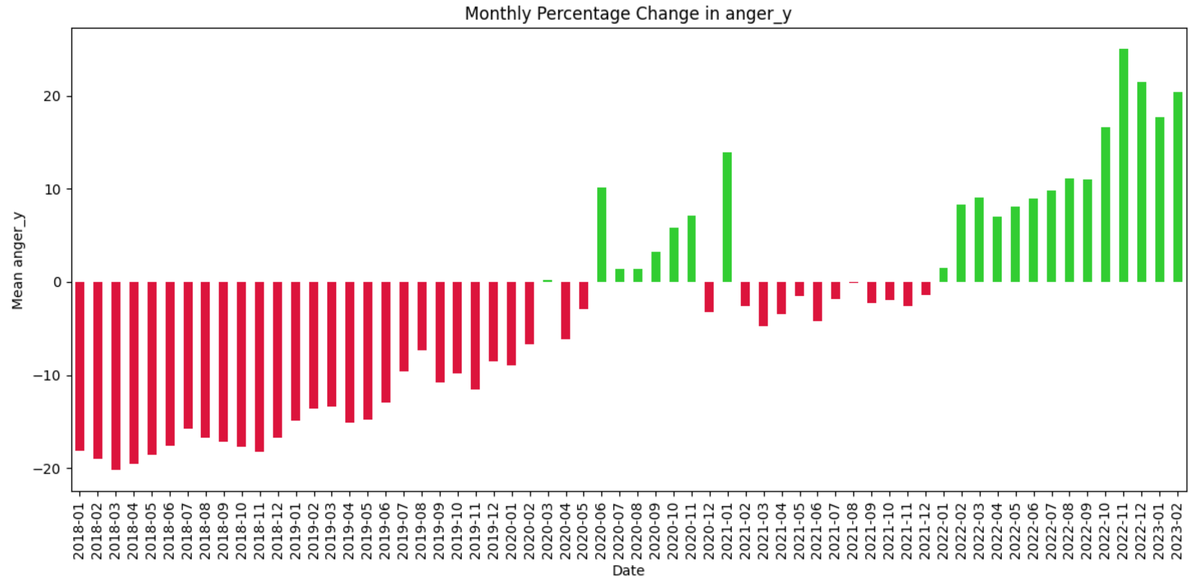


The graph depicting the monthly percentage change in fear across various topics illustrates that most areas were significantly impacted by COVID-19. While all topics showed an increase in fear during the early months of the pandemic and again during the summer/fall of 2021, Science and Technology seems to be less affected compared to others, which might be due to the nature of the topic.

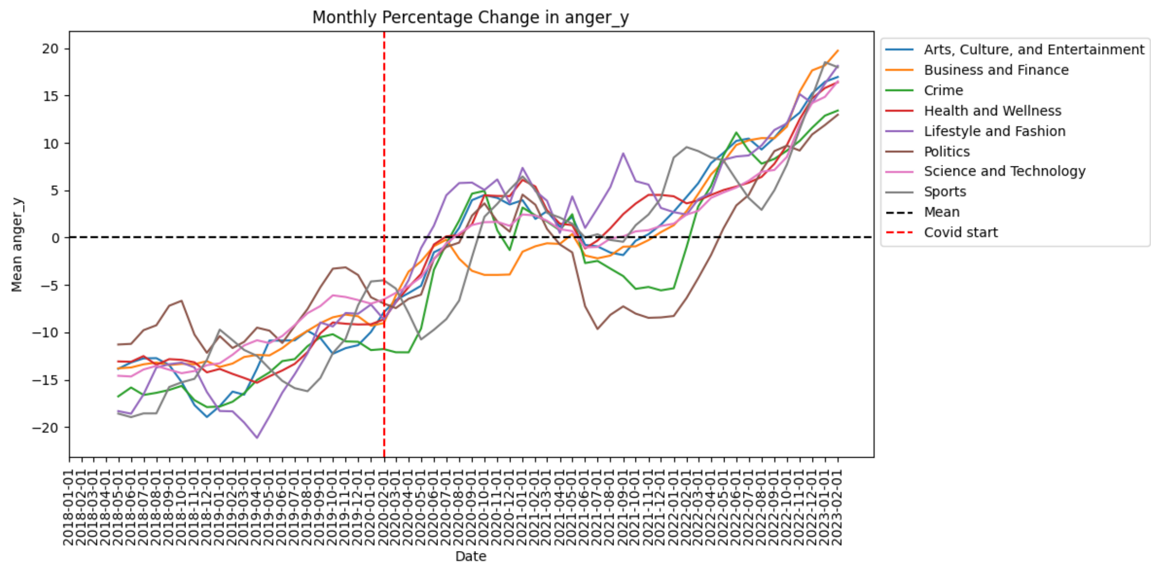
Leisure-related topics such as Arts, Sports, and Lifestyle experienced a resurgence in fear later into the pandemic, particularly towards the end of 2022. This could be linked to ongoing uncertainties in these sectors due to the new COVID-19 waves and related restrictions.

Business and Finance, followed by Health and Wellness, are the topics that registered the highest spikes in fear. The fear experienced during the COVID-19 pandemic can be attributed to financial instability, unpredictable market fluctuations, unemployment, and health concerns directly related to COVID-19. These factors shed light on how various facets of life were impacted by fear during the pandemic, reflecting wider societal concerns.

ANGER



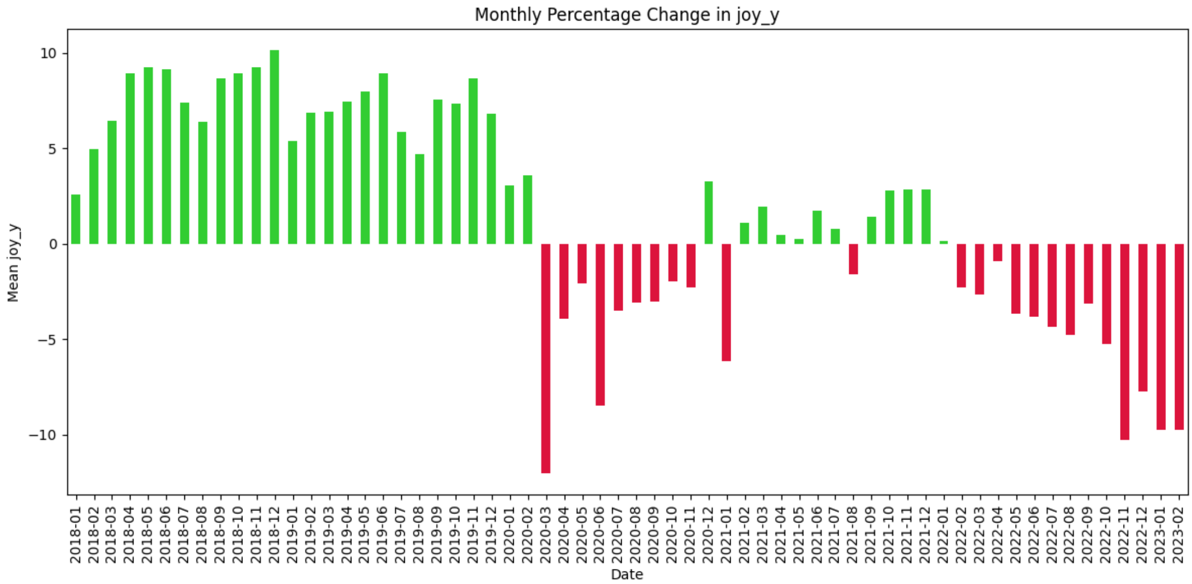
The analysis of anger trends on Twitter during the COVID-19 pandemic reveals an escalating pattern that intensifies with major global and political events. Starting with a spike at the onset of the pandemic, anger increases dramatically again in June 2020, coinciding the global protests following George Floyd's death. Another significant increase in January 2021 aligns with the US presidential elections and the Capitol assault. Furthermore, anger levels go upward starting from February 2022, notably due to Russia's invasion of Ukraine.



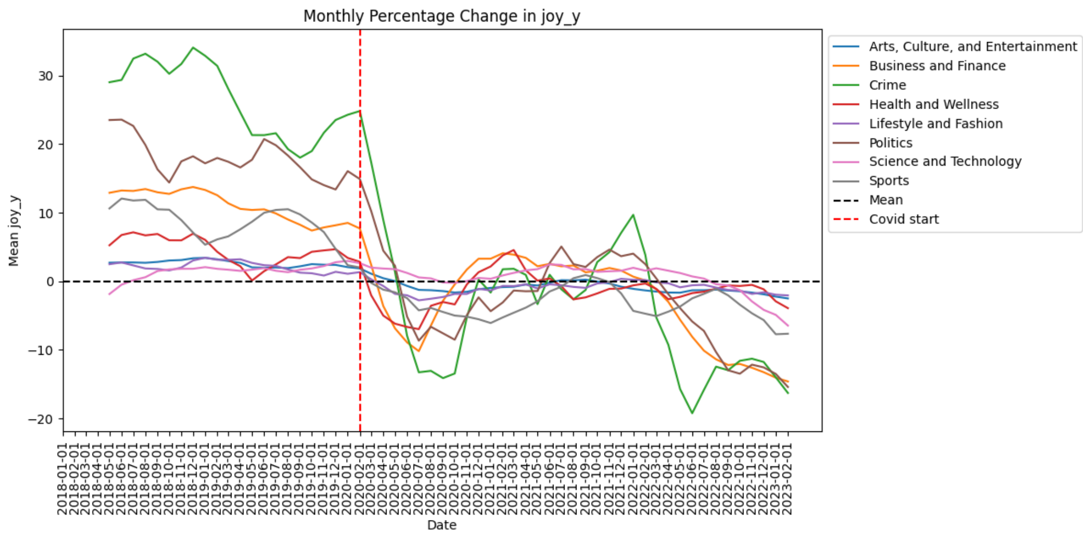
The analysis of anger across all topics on Twitter reveals a general increase in anger over time. However, during the summer and fall of 2021, some topics experienced a noticeable decrease in anger levels. This dip could be attributed to positive developments such as the vaccine rollout and easing of restrictions, which may have temporarily alleviated public frustrations.

Among the topics analyzed, Politics exhibited the most significant fluctuations in anger. This variance is likely due to the topic's sensitivity to political events, such as elections and significant policy changes, which tend to evoke strong emotional responses.

JOY



Initially, before the onset of the pandemic, joy is at high levels, but as the pandemic unfolds, there's a drastic decline. The low levels of joy continue until December 2020, when several countries authorized vaccines like Pfizer and Moderna, which promised to help combat the pandemic. However, this positive trend is briefly disrupted in January 2021 due to the Capitol Building assault in the US. Following these events, joy begins to recover again, only to face another disruption by the onset of the invasion of Ukraine.



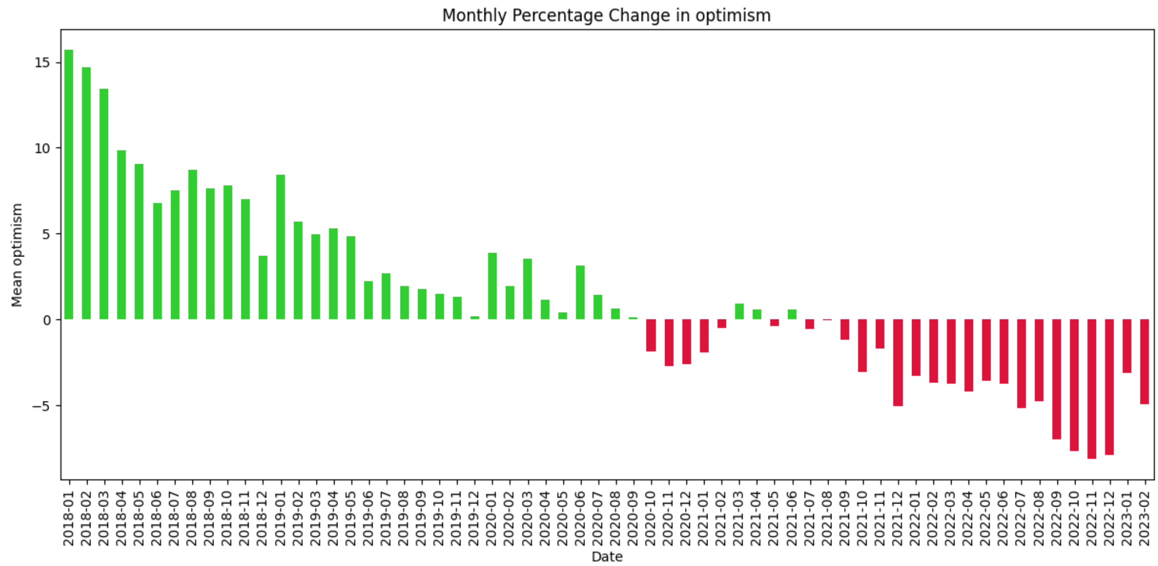
The monthly joy percentage change graph during the COVID-19 pandemic reveals distinct emotional impacts across different areas.

With the beginning of the pandemic, a decline in joy is noticeable across most topics. Interestingly, the topics of Science and Technology show resilience during this period, remaining stable. Similarly, Arts and Culture, along with Lifestyle and Fashion, experience only a slight decrease in the joy index, perhaps due to the solace they offer during times of crisis.

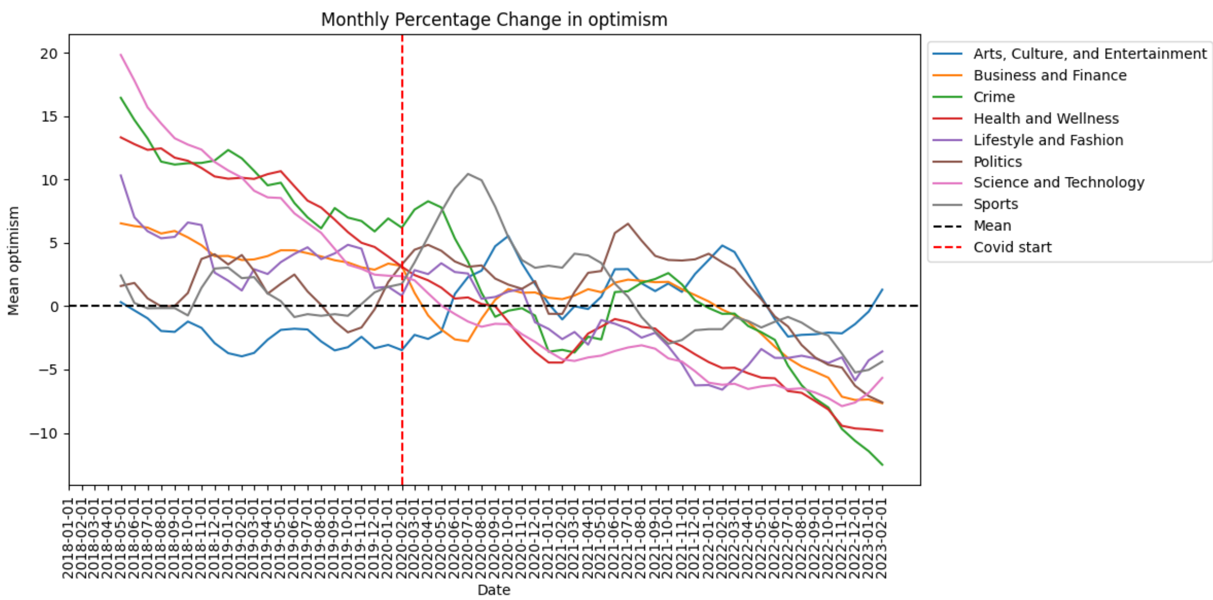
In stark contrast, the topics of Crime, Politics, and Business and Finance show significant drops in joy. These areas were directly impacted by the pandemic's broader effects—increased crime rates due to economic desperation, political unrest from government handling of the pandemic, and financial instability affecting businesses and economies worldwide. The significant decline in joy within these topics highlights the negative impact resulting from the pandemic's socioeconomic challenges.

This analysis shows how the pandemic affected people's feelings differently in different parts of life. While Science and Technology could keep people hopeful with good news, Politics and Business took a hit, showing how various factors can affect public mood during crises.

OPTIMISM



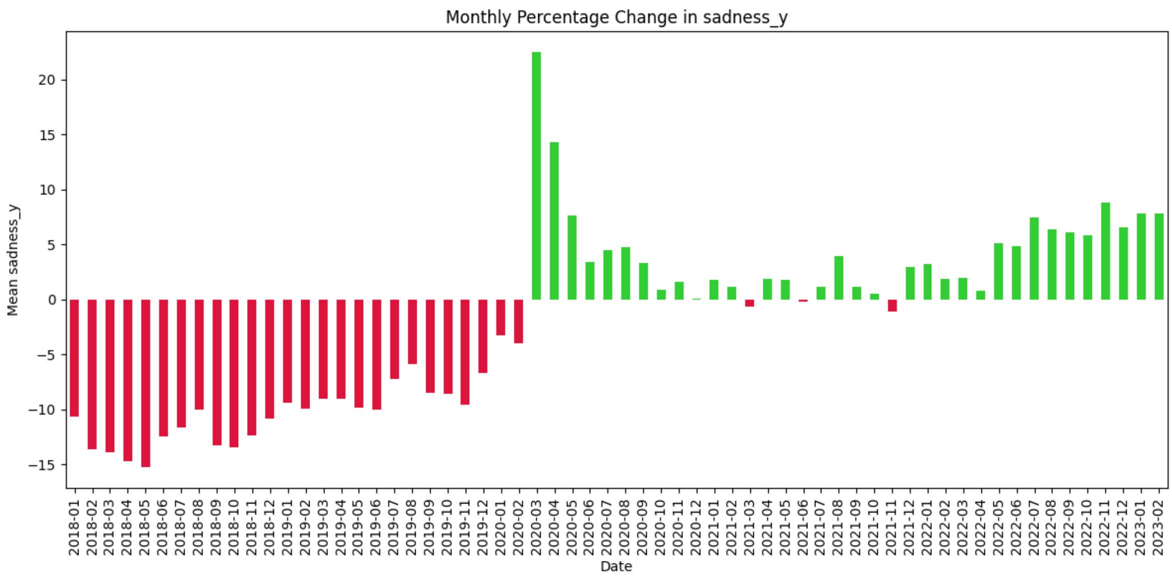
The graph of monthly optimism change during COVID-19 suggests that optimism didn't immediately drop when the pandemic began. This might mean people initially remained hopeful, however as challenges such as new waves and variants emerge, a clear decline in optimism becomes evident. This trend intensifies with significant global events, notably the war in Ukraine and the acquisition of Twitter by Elon Musk.



During the COVID-19 pandemic, the impact on optimism varied significantly across different topics. In the areas of Science and Technology, Health and Wellness, and Crime, a steady decline in optimism was evident. On the other hand, Arts, Culture, and Entertainment showed relative stability in optimism, suggesting that these sectors provided some escapism to the public, helping them maintain a more optimistic outlook.

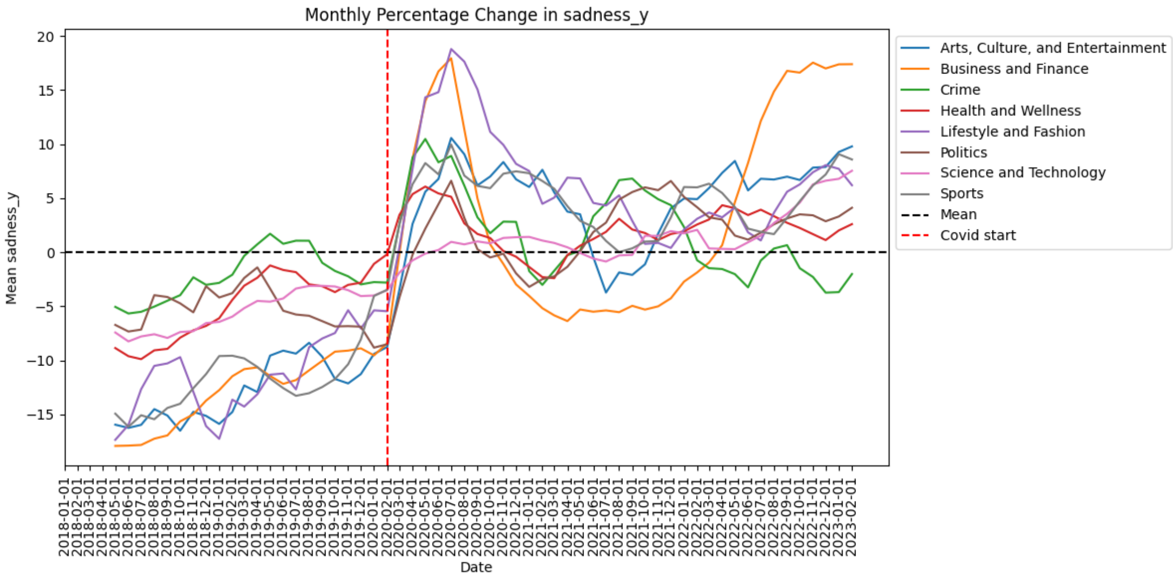
The Business and Finance sector was affected quickly by COVID-19, reflecting immediate economic uncertainties. In contrast, the Sports sector initially saw an increase in optimism despite major disruptions such as the cancellation of the 2020 Olympics. Optimism briefly surged again during the summer of 2021 with the holding of the Tokyo Olympics, underscoring how major events can positively influence sentiment within specific topics.

SADNESS



The graph depicting the monthly percentage change in sadness during the COVID-19 pandemic highlights an immediate and sharp increase in sadness as the pandemic begins, reflecting the global distress at the time. Following this initial spike, there appears to be a stabilization in sadness levels, from the end of 2020 to the middle of 2022, suggesting a period of adaptation or returning to the new normal.

However, this relative stability in sadness is disrupted by another significant increase towards the middle of 2022 that stays persistent into the following year.

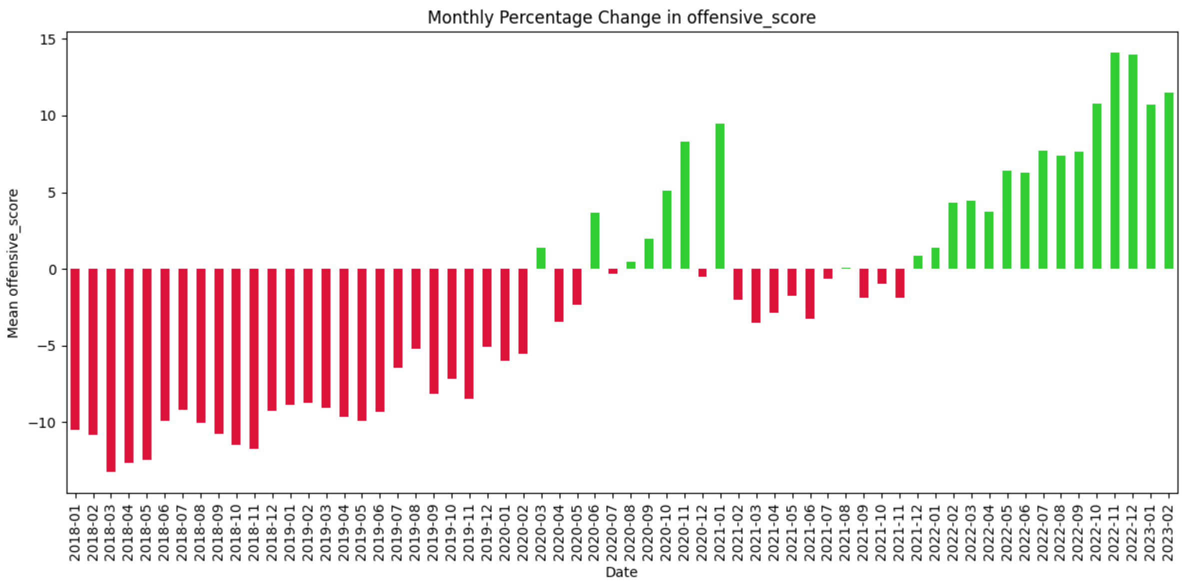


The graph showcasing the monthly percentage change in sadness across various topics during the COVID-19 pandemic reveals how different sectors reacted emotionally during key periods of the crisis. At the start of the pandemic, sadness rose universally across all topics, with the most impacted ones being Business and Finance and Lifestyle and Fashion. Both of these topics' sadness index however subsequently decreased as restrictions were eased and things started returning back to normal in 2021, with the introduction of COVID-19 vaccination policies.

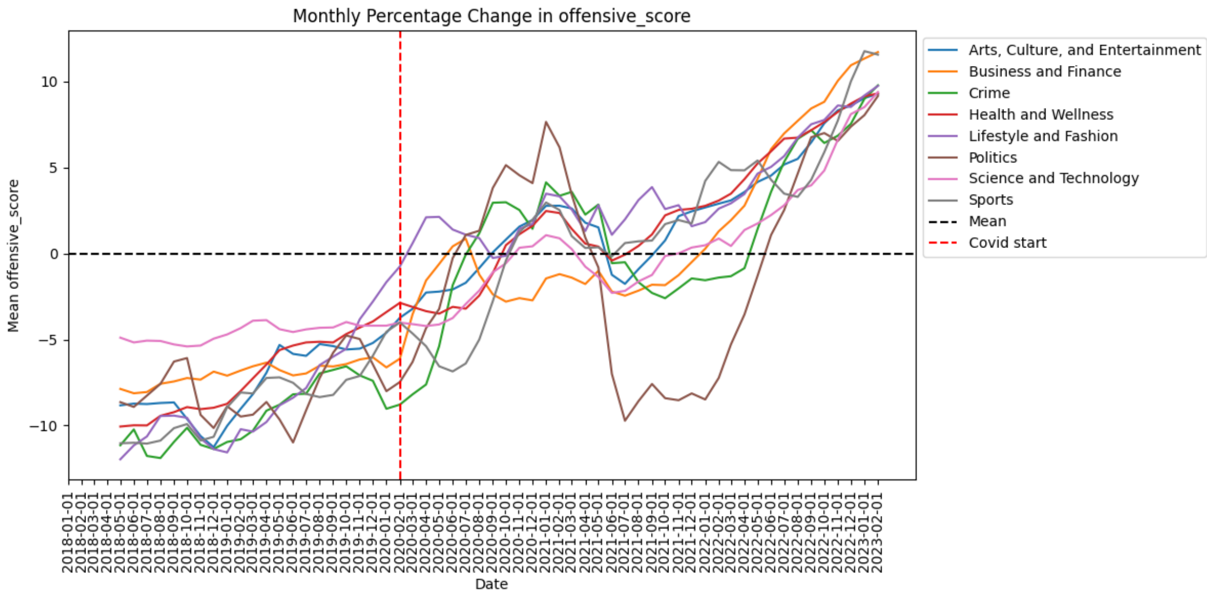
Interestingly, early to mid-2022 shows another notable rise in sadness within the Business and Finance topic which can be linked to the economic repercussions following the outbreak of the war in Ukraine. This conflict caused widespread economic effects, such as disruptions to global supply chains, higher energy prices, and increased instability in financial markets. The sadness expressed in tweets from this sector probably reflects worries about these destabilizing impacts.

Meanwhile, the topic of Science and Technology displayed a distinct pattern with a steady increase in sadness throughout the entire period studied. This consistent rise might reflect the continuous pressure and high expectations placed on this sector to innovate and provide solutions during the pandemic.

OFFENSIVE SCORE



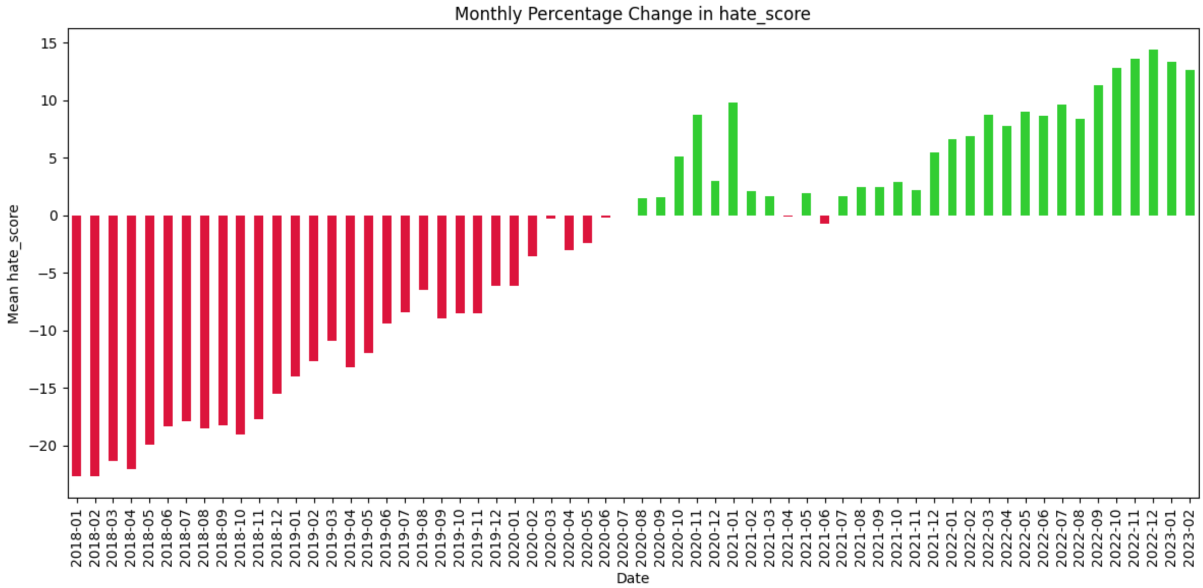
The offensiveness graph shows consistent negative values before the pandemic, but it spikes at its onset, reflecting heightened tensions. Significant surges in offensive language correlate with major events such as the Black Lives Matter protests, renewed COVID-19 lockdowns between August and October 2020, the Capitol Building storming in January 2021, and the early 2022 invasion of Ukraine. Each of these events marks a clear escalation in offensiveness, showcasing how societal unrest and major global crises contribute to increased aggressive and offensive language.



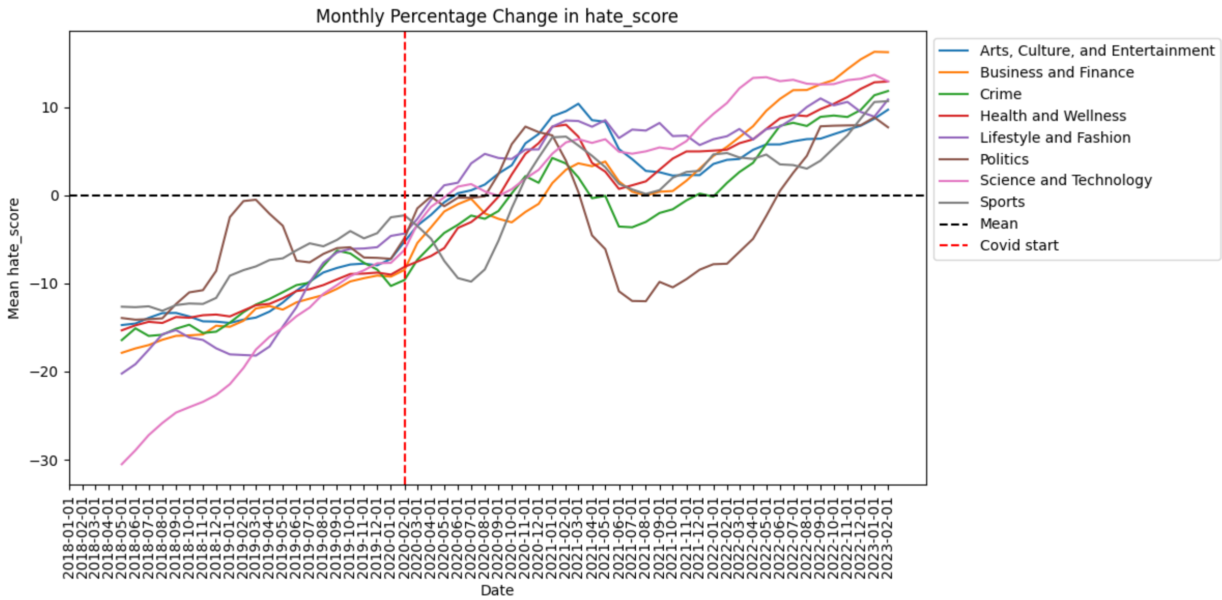
The analysis of the offensive score by topic reveals a general increase in offensiveness across all topics over time on Twitter. Most topics show a significant peak in offensiveness in late 2020, aligning with the announcement of COVID-19 vaccines, which may have triggered debates and polarized opinions, contributing to a rise in offensive language. An exception is noted in the Business and Finance sector, which reaches its peak earlier, in the summer of 2020, possibly due to intense economic pressures and uncertainties during the initial stages of the pandemic.

Political tweets consistently exhibit the most turbulence in terms of offensiveness. Notably, there are spikes during charged events such as the US election and the storming of the Capitol Building, where there were increased political tensions. Following these events, there was a noticeable decline in offensiveness, suggesting a temporary easing of tensions. However, offensiveness in political discussions escalates again with the onset of the invasion of Ukraine, indicating that geopolitical conflicts provoke aggressive communications on social media.

HATE



The graph illustrating the monthly percentage change in hate score over time during the COVID-19 pandemic reveals an important shift from below-average levels to above-average levels, highlighting a gradual but consistent increase in hate speech. The initial significant uptick in hate occurs around August 2020, intensifying by October 2020. This period corresponds with a resurgence of COVID-19 cases worldwide, which led to renewed lockdowns and restrictions. Following this period, the rise in hate maintains a steady pace until it escalates further with the invasion of Ukraine, marked by a continuous monthly increase afterwards. This pattern underscores how prolonged crises and significant geopolitical events can fuel a steady rise in hate speech across social platforms.

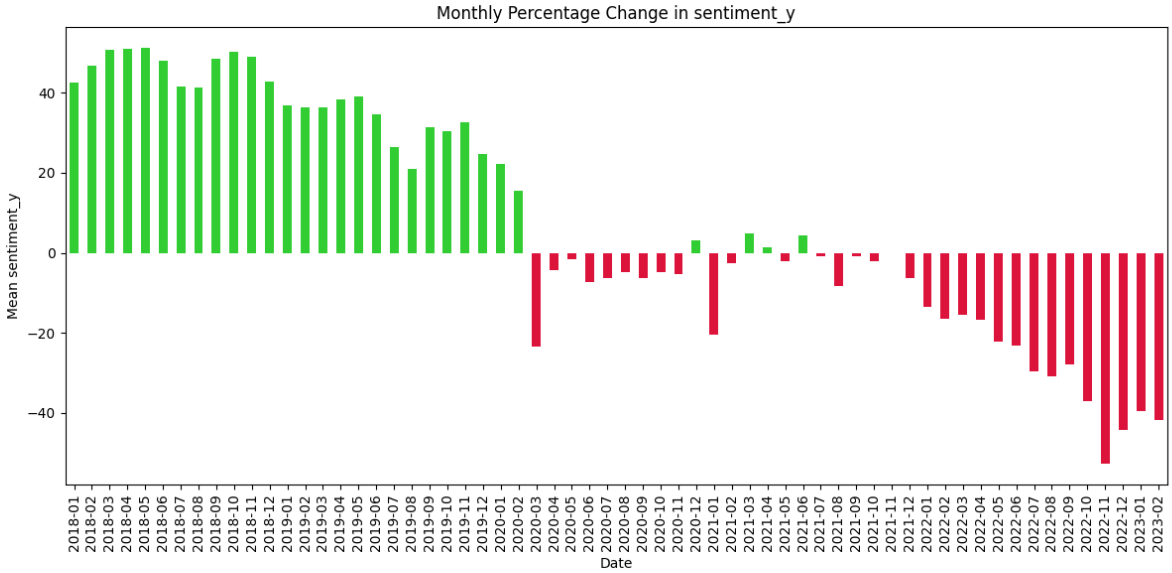


The topic graph indicates a consistent pattern where hate gradually increases over time for all topics. This trend is reflective of the growing unrest and tension during the COVID-19 pandemic and other subsequent global crises.

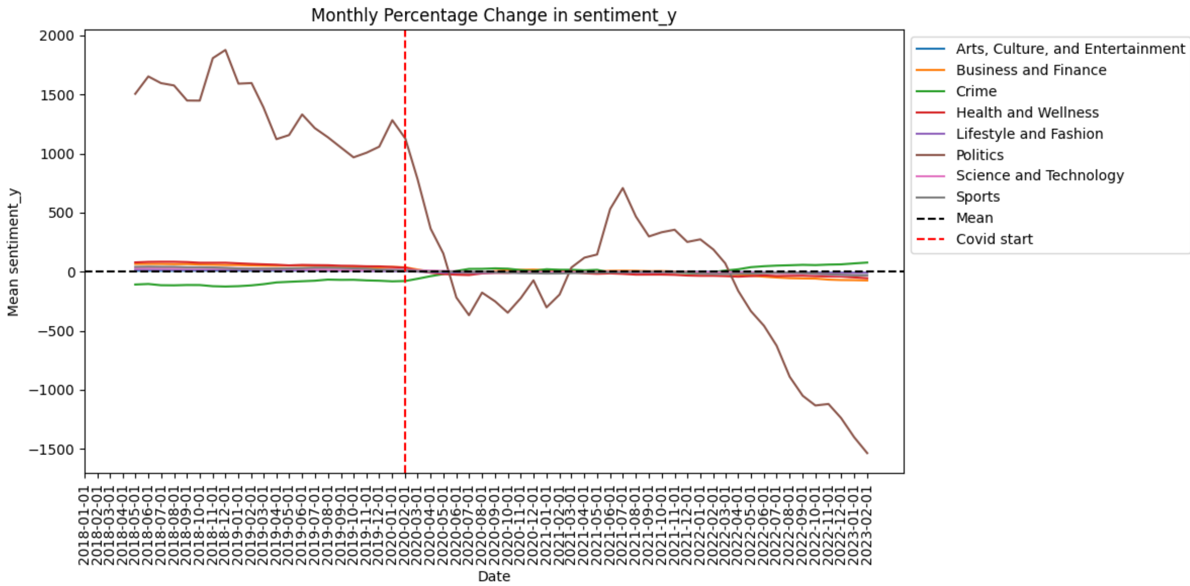
Interestingly, a temporary decline in hate is observed for some topics during the summer and fall of 2021. This drop might be associated with positive developments in the pandemic situation, such as the ongoing vaccination campaigns which led to a reduction in COVID-19 case numbers.

The topic of Politics, however, shows the most significant fluctuations in hate, with marked peaks around major political events such as elections.

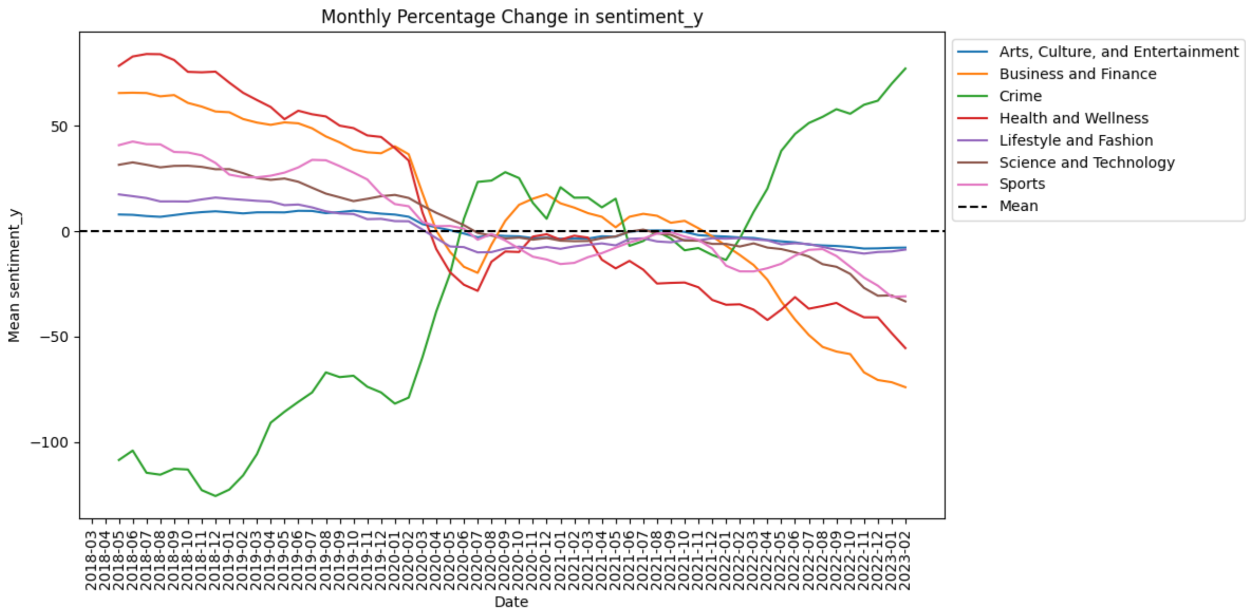
SENTIMENT



The sentiment graph illustrates a clear trend where sentiment was above average prior to the pandemic but shifted significantly as the pandemic unfolded. Initial spikes in negative sentiment are evident at the beginning of the pandemic, reflecting the global shock and uncertainty. Another notable spike occurs during the storming of the Capitol Building, highlighting a moment of intense political turmoil. Following the invasion of Ukraine, a more significant and sustained decrease in sentiment is observed, indicating a deep and lasting impact on global emotions.



A huge variation was noted in sentiment when it came to Political tweets. A sharp decline in the sentiment index is noted in the early months of COVID, and although there was some improvement by mid-2021, the overall trend in tweets about Politics has been generally downward. This negative trajectory culminated in an all-time low in February 2023, the most recent data point in our study.



We have now set aside Political tweets due to their extreme variability, in order to see the trends for the other emotions better. Most topics demonstrate a declining sentiment trend over time, particularly pronounced in sectors like Business and Finance, and Health and Wellness. These areas were among the hardest hit at the beginning of the pandemic. In the case of Business and Finance, the immediate financial instability, job losses, and market crashes drove sentiment downwards. Similarly, Health and Wellness faced challenges due to the sudden strain on healthcare systems, concerns over public health safety, and the direct impacts of the virus.

In stark contrast, the sentiment in Crime-related tweets increased during these periods. This rise during global crises like the COVID-19 spread and the Ukraine war may seem counterintuitive given the typically negative connotations associated with crime. The rise in positive sentiment for Crime-related tweets might be because people are talking more about how communities stick together, the important role of the police in keeping things orderly, and a general agreement against crime during tough times.

In conclusion, the analysis of emotional responses across various topics during significant global crises, particularly the COVID-19 pandemic, reveals diverse impacts on public sentiment. Sectors like Business and Finance, and Health and Wellness experienced heightened fear and sadness due to financial instability and health concerns, underscoring the severity of the pandemic's impact. Conversely, Science and Technology maintained a relative level of stability, reflecting the hope brought about by technological advances and medical breakthroughs. Political discourse displayed pronounced spikes in anger and hate, driven by controversial political events.

These differing trends highlight how specific sectors can evoke distinct public emotions and discussions, influenced heavily by their direct involvement and impact during global events.

CHAPTER 5: DISCUSSION

5.1 KEY FINDINGS

The findings from our analysis of emotional fluctuations during the COVID-19 pandemic highlight significant trends in tweets across global and local contexts, as well as across various topics. In this section we are providing a summary of our most important findings.

In the worldwide analysis of sentiment correlations during the COVID-19 pandemic, key findings reveal interconnected dynamics between emotions and policy responses. Positive emotions consistently showed strong correlations with other positive emotions, a pattern mirrored among negative emotions as well. Positive emotion indices were generally higher than those of negative emotions.

Regarding COVID-19 restrictions, measures such as workplace closures, school closures, and stay-home requirements were typically implemented together. In contrast, facial coverings were often adopted independently of these broader restrictions. There was a noticeable negative relationship between the extent of COVID-19 restrictions and the availability of vaccinations, suggesting that increased vaccine accessibility could lead to relaxation of certain measures.

Interestingly, the analysis did not demonstrate strong relationships between the severity of the pandemic and the stringency of restrictions, likely because many of the strictest measures were introduced early in the pandemic, when the actual numbers of cases and deaths were still comparatively low, suggesting a proactive rather than reactive approach to public health policy.

The case study of emotional fluctuations in the USA during the COVID-19 pandemic highlighted significant initial spikes in negative emotions like fear, and sadness, and drops in positive emotions like joy pointing to a sharp, reactive public sentiment at the start of the crisis. The anger, offensive score and hate metrics were also elevated early in the pandemic and escalated further during subsequent global events, particularly during periods of extended lockdowns. Optimism appeared largely unimpacted directly by COVID-related developments but followed a general downward trend, suggesting a slow erosion of public positivity over time.

T-test analyses revealed a clear shift towards negativity with overall sentiment being the most impacted. They also shed light on the different nature of emotions. The surge in fear and sadness early on in the pandemic highlights the innate human response to abrupt changes caused by the crisis. Conversely, metrics such as hate, offensiveness, and optimism were more resilient to drastic changes during the first months of pandemic but showed significant shifts later on, suggesting that prolonged uncertainty fueled their escalation. Other emotions like anger and hate appeared more responsive to ongoing political and social events rather than the pandemic.

From the fixed effects regressions, there was a noticeable trend of increased hostility in social media interactions, with both hate scores and offensiveness rising on an individual level, while joy saw a significant reduction. This trend underscores a major shift in the nature of online interactions, reflecting a growing polarization within the community as the pandemic progressed. This analysis highlights the complexities of immediate and long-term reactions to pandemic-related events.

In the Swedish case study, the emotional responses to the COVID-19 pandemic presented distinct patterns compared to other nations, notably influenced by Sweden's unique approach to managing the crisis.

The sentiment index demonstrated a decline with the onset of COVID-19, accompanied by sharp increases in fear and sadness, while hate metrics showed a general upward trajectory. Meanwhile joy significantly dropped at the beginning and never returned to pre-pandemic levels, showing the lasting impact of COVID-19 on positive emotions. Interestingly, anger and offensive scores were not impacted by COVID-19, which was attributed to Sweden's less restrictive pandemic measures. Unlike other positive emotions optimism showed no significant drop during the pandemic itself and showed a notable increase with the start of Sweden’s vaccination campaign, underscoring the positive impact of significant health-related developments on public sentiment.

T-test analyses indicated a general trend towards more negative emotions such as anger, sadness, and hate across the span from 2018 to 2023, with positive emotions like joy and overall sentiment declining. This points to a gradual but persistent shift towards a more negative emotional landscape despite the lighter restrictions.

Fixed effects regressions revealed that while internal emotional experiences like joy, sadness, and optimism were significantly affected by the pandemic on an individual level, there was no substantial change in the outward expressions of fear, anger, or offensive language among individual users. This pattern suggests that Swedish Twitter users might have experienced internal shifts in emotion due to the pandemic but showed a lesser tendency to express these emotions outwardly. This aligns with a cultural tendency in Sweden to manage emotions internally rather than displaying them in forms of public aggression or hostility. This nuanced analysis of emotional expression highlights the complex interplay between cultural norms, individual behavior, and public health crises.

In comparing the emotional responses to the pandemic between Sweden and the USA, distinct differences emerge, reflecting the influence of each country's social, political, and cultural contexts. In Sweden, the t-tests revealed minimal changes in hate and offensive behaviors, while the USA experienced significant increases in both emotions. This implies that the atmosphere in the USA might have been more supportive of such manifestations during the pandemic.

The differences became more prevalent when comparing the fixed effects analysis. In the USA, the pandemic had substantial impact on individuals, with significant rises in negative emotions and declines in positive emotions and overall sentiment. In contrast, Sweden showed no statistically significant changes in fear, anger, or sentiment. And while, hate and offensive scores increased significantly in the USA, Sweden exhibited changes only in hate and not offensiveness. These findings underscore potential differences in emotional responses to the crisis between the two countries, shaped by social, and governmental factors.

The topic modeling analysis of emotional responses during the pandemic revealed several important trends across different subjects. Fear experienced a universal surge at the pandemic's early days, with Business & Finance and Health & Wellness topics showing the most pronounced spikes, while Science & Technology displayed relative resilience. Anger exhibited a general upward trend across all topics, with Politics displaying the most drastic fluctuations, indicative of the nature of political discourse during crises.

Joy declined across all topics, with Science and Technology remaining relatively stable and offering hope through positive news, while Politics and Business & Finance suffered significant drops. Surprisingly, optimism did not decline significantly with the spread of COVID-19, but sectors like Science & Technology and Health & Wellness saw a steady decrease over time, contrasting with relatively stable optimism levels in Arts, Culture, & Entertainment.

Sadness levels were heightened in Business & Finance and Lifestyle & Fashion topics, normalizing as pandemic related restrictions were lessened. Meanwhile Science & Technology showed a steady increase in sadness over time. Offensive and hate scores gradually increased across all topics with the pandemic, with political tweets being the most turbulent.

Overall, sentiment varied widely across different contexts, with political tweets displaying significant fluctuations, and pandemic affected sectors like Business & Finance and Health & Wellness showing a decline over time.

These insights collectively present the ways in which different communities and sectors responded to the global crisis. First, the research underscores that public sentiment is highly dynamic and responsive to new developments. As seen throughout this study, emotions fluctuated in response to events such as government policies, vaccine rollouts, and other social and political phenomena.

Additionally, the comparison between the USA and Sweden showed how government policies critically influence public sentiment during crises. Different national approaches led to different emotional landscapes, demonstrating the profound impact of policy decisions on societal reactions. Furthermore, the study emphasized the pivotal role of social media in crisis communication. It showed that social media platforms play a crucial role in reflecting public sentiment in real time.

5.2 CONTRIBUTION

1. LONG-TERM IMPACT OF COVID-19 ON PUBLIC SENTIMENT

This study explores the emotional impact of COVID-19 on public sentiment through a detailed analysis spanning tweets from January 2018 to February 2023. This extended period allows for an in-depth observation of both pre-pandemic emotional states and their fluctuations throughout the duration of the COVID-19 pandemic and afterwards. The span of the data allowed us to examine both short-term and long-term effects, providing a comprehensive understanding of how public sentiment evolved during the crisis. This sheds light on the cumulative psychological toll of the pandemic and related measures on public well-being.

2. ANALYSIS ON SENTIMENT FLUCTUATION RELATED TO SPECIFIC EVENTS

When explaining the observed fluctuations in emotions, we took into account both COVID-19 related and important unrelated social and political events. By analyzing those events and their immediate impacts on public sentiment, we aimed to distinguish their effects from that of the pandemic. This approach is essential for accurately identifying the factors driving emotional fluctuations over an extended period. For instance, the analysis considered the influence of major political elections, social movements, and international conflicts, providing explanations for shifts in public mood that may otherwise be inaccurately attributed solely to the pandemic. In that way, our sentiment analysis allows for the complex relationship between different types of events and the public's emotional response to emerge.

3. COMPARATIVE ANALYSIS OF SENTIMENT BETWEEN DIFFERENT NATIONAL STRATEGIES

This thesis offers a unique perspective by comparing the sentiments across different national strategies to COVID-19, presenting how governmental approaches impact public emotions. By examining the emotional responses in countries with distinct pandemic management strategies, such as the stringent measures adopted in the USA versus the more voluntary approach taken by Sweden, this study uncovers the significant role that national policies play in shaping public emotions during a crisis.

This analysis takes a cross-national approach, which is different from most studies commonly found in existing literature. It shows how government policies can affect how people feel, demonstrating that the nature and strictness of health measures can result in vastly different emotional responses. This can help us make better choices in the future when it comes to crisis management.

4.INTEGRATION OF MULTIPLE SENTIMENT ANALYSIS TOOLS IN REAL-TIME EVENT IMPACT ANALYSIS

Multiple sentiment analysis tools were used to assess the impact of events on public sentiment. Both traditional lexicon-based methods and advanced machine learning techniques, including VADER, TextBlob, and RoBERTa were utilized. The combination these methodologies enables a more robust and multifaceted analysis of sentiment, allowing us to interpret complex emotional expressions.

By employing multiple tools, this thesis addresses the limitations that single-method studies might encounter, such as biases inherent in specific analytical models or their sensitivity to different types of language use, such as slang or informal online communication.

5.3 FUTURE WORK

EXPAND ON CROSS-CULTURAL SENTIMENT ANALYSIS

Expanding the comparative analysis to include more countries with diverse cultural backgrounds and different approaches to crisis management can significantly enhance our understanding of global sentiments during crises. Additionally, by incorporating sentiment analysis across multiple languages and cultural contexts would allow for a more spherical understanding of the sentiment undertones present in non-English speaking countries. This approach would allow for a better examination of regional differences in emotional reactions and the effectiveness of policy measures. This can help international organizations and governments tailor their policies to be culturally sensitive and effective across different global communities.

ANALYSIS OF SENTIMENT RESILIENCE

Investigating how quickly public sentiment rebounds following crises or major events can provide insights into the resilience of communities. This research extension could focus on the "recovery" phase, assessing how different demographic groups regain a sense of normalcy or optimism following an event that caused a significant emotional shift. By measuring the time needed for emotional recovery across different populations, we can identify factors that contribute to resilience and vulnerability. This could help create mechanisms that accelerate recovery and reinforce resilience in future crises.

IN DEPTH TOPIC-SPECIFIC SENTIMENT ANALYSIS

Extending upon the sentiment analysis on specific topics such as healthcare, business, lifestyle etc. can output targeted insights that are crucial for sector-specific decision-making.

For example, analyzing sentiment within the healthcare sector during a health crisis can inform healthcare providers about the concerns of healthcare workers and patients. Similarly, understanding the sentiment in the business sector during stay-at-home mandates can help executives make informed decisions about remote work strategies.

EFFECTIVENESS OF POLICY INTERVENTIONS ON SENTIMENT

Conducting detailed studies to measure the effectiveness of specific policy interventions on public sentiment can provide valuable feedback on governmental actions. For instance, analyzing public sentiment before and after the implementation of lockdowns, vaccination campaigns, or economic relief measures can offer insights into how these policies are perceived by the public.

REFERENCES

1. IBM, "What Is Sentiment Analysis?," IBM, Accessed: May 08, 2024. [Online]. Available: https://www.ibm.com/cloud/learn/sentiment-analysis
2. World Health Organization, "Impact of COVID-19 on people's livelihoods, their health and our food systems," Oct. 13, 2020. [Online]. Available: https://www.who.int/news/item/13-10-2020-impact-of-covid-19-on-people's-livelihoods-their-health-and-our-food-systems
3. S. Boon-itt and Y. Skunkan, "Public Perception of the COVID-19 Pandemic on Twitter: Sentiment Analysis and Topic Modeling Study," JMIR PUBLIC HEALTH AND SURVEILLANCE, vol. 6, e21978, 2020. [Online]. Available: https://doi.org/10.2196/21978
4. R. Chandrasekaran, V. Mehta, T. Valkunde, and E. Moustakas, "Topics, Trends, and Sentiments of Tweets About the COVID-19 Pandemic: Temporal Infoveillance Study," Journal of Medical Internet Research, vol. 22, e22624, 2020. [Online]. Available: https://doi.org/10.2196/22624
5. C. Biliotti, F. J. Bargagli-Stoffi, N. Fraccaroli, M. Puliga, and M. Riccaboni, "Breaking Down the Lockdown: The Causal Effects of Stay-At-Home Mandates on Uncertainty and Sentiments During the COVID-19 Pandemic," May 2023.
6. J. Lyu, E. Han, and G. Luli, "Topics and Sentiments in COVID-19 Vaccine-related Discussion on Twitter (Preprint)," Journal of Medical Internet Research, vol. 23, 2021. [Online]. Available: https://doi.org/10.2196/24435
7. B. S. Ainapure, R. N. Pise, P. Reddy, B. Appasani, A. Srinivasulu, M. S. Khan, and N. Bizon, "Sentiment Analysis of COVID-19 Tweets Using Deep Learning and Lexicon-Based Approaches," Sustainability, vol. 15, no. 3, p. 2573, 2023. [Online]. Available: https://doi.org/10.3390/su15032573
8. M. Mushtaq, M. S. Fareed, M. Almutairi, S. Ullah, G. Ahmed, and K. Munir, "Analyses of Public Attention and Sentiments towards Different COVID-19 Vaccines Using Data Mining Techniques," Vaccines, vol. 10, 661, 2022. [Online]. Available: https://doi.org/10.3390/vaccines10050661

9. Arora, K. Chopra, D. Chaudhary, I. Gorton, and B. Kumar, "Sentiment Analysis of Social Media Data on COVID-19," in PROC. COMPUTER SCIENCE & INFORMATION TECHNOLOGY (CS & IT), D. C. Wyld et al., Eds., 2023, pp. 13-28.
10. Q. Zhang, G. Y. Yi, L. P. Chen, and W. He, "Sentiment analysis and causal learning of COVID-19 tweets prior to the rollout of vaccines," PLoS One, vol. 18, no. 2, e0277878, Feb. 2023. [Online]. Available: https://doi.org/10.1371/journal.pone.0277878. PMID: 36827382; PMCID: PMC9955611.
11. Mehta, H., Ivanovic, J., Tomic, I., & Nikolov, N. S. (2020). Characteristics of Registered Clinical Trials Assessing Treatments for COVID-19: A Cross-Sectional Analysis.
12. Milani, M. (2020). COVID-19 Outbreak Response: Effects Analysis Cross-Country.
13. Gadarian, S. K., Goodman, S. W., & Pepinsky, T. B. (2020). Partisanship, Health Behavior, and Policy Attitudes in the Early Stages of the COVID-19 Pandemic.
14. Hartnett, K. P., Kite-Powell, A., DeVies, J., Coletta, M. A., Boehmer, T. K., Adjemian, J., Gundlapalli, A. V., & National Syndromic Surveillance Program Community of Practice. (2020). Impact of the COVID-19 Pandemic on Emergency Department Visits - United States, January 1, 2019-May 30, 2020.
15. Lange, S. J., Ritchey, M. D., Goodman, A. B., Dias, T., Twentyman, E., Fuld, J., Schieve, L. A., Imperatore, G., Benoit, S. R., Kite-Powell, A., Stein, Z., Peacock, G., Dietz, E., & Boehmer, T. K. (2020). Indirect Effects of the COVID-19 Pandemic on Emergency Department Visits - United States, January-May 2020.
16. Bento, A. I., Nguyen, T., Wing, C., Lozano-Rojas, F., Ahn, Y.-Y., & Simon, K. (2020). Evidence from Internet Search Data Shows Information-Seeking Responses to News of Local COVID-19 Cases.
17. Hu, Y., Pan, S., Zhou, C., & Zhou, D. (2021). Revealing Opinion towards COVID-19 Vaccines with Twitter
18. Ahaotu, J. O. (2023). A Multimodal Discourse Analysis of Selected Social Media Posts on the #BlackLivesMatter Protest. Journal of Discourse Studies. DOI: 10.32996/jpda.2023.2.1.4

19. D. J. Hopkins, "The American Election 2020: Contexts and consequences—the consequences of elections in pandemic times," PS: POLITICAL SCIENCE & POLITICS, vol. 54, no. 1, pp. 1-6, Jan. 2021.
20. J. Goldsmith, "The Capitol riot: Anomaly or tipping point?" LAWFARE, Jan. 2021.
21. Goldthau and L. Hughes, "The consequences of the war in Ukraine for the geopolitics of energy," ENERGY RES. SOC. SCI., vol. 91, Article 102405, May 2022.
22. Pan, H. (2023). Analysis of Motivations, Process, and Implications of Elon Musk's Acquisition of Twitter. DOI: 10.54691/bcpbm.v47i.5185
23. Oh, H., Goh, K. Y., & Phan, T. Q. (2018). Are You What You Tweet? The Impact of Sentiment on Digital News Consumption and Social Media Sharing. SOCIAL SCIENCE RESEARCH NETWORK. DOI: 10.2139/ssrn.3215634. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3215634
24. Chen, R. (2022). "A timeline of events," in A TIMELINE OF EVENTS, Norwegian University of Science and Technology, Trondheim, Norway, Dec. 2022. [Online]. Available: https://doi.org/10.4324/9781003289364-3
25. N. Aylott and N. Bolin, "A party system in flux: the Swedish parliamentary election of September 2018," WEST EUROPEAN POLITICS, vol. 42, no. 6, pp. 1504-1515, 2019. [Online]. Available: https://consensus.app/papers/party-system-election-september-2018-aylott/
26. S. Boulianne, M. Lalancette, and D. Ilkiw, "'School Strike 4 Climate': Social Media and the International Youth Protest on Climate Change," MEDIA AND COMMUNICATION, vol. 8, no. 2, pp. 208-218, 2020. [Online]. Available: https://consensus.app/papers/school-strike-climate-social-media-international-youth-boulianne/b7d98c4e676b56779ddeea96b3c3e94b/.
27. "Sweden Joins NATO as War in Ukraine Prompts Security Rethink," TODAY, Mar. 7, 2024. [Online]. Available: https://www.todayonline.com/sweden-joins-nato-as-war-in-ukraine-prompts-security-rethink
28. Schramm-Nielsen, J. (2002). Conflict Management in Scandinavia. Retrieved from Conflict Management in Scandinavia
29. "Twitter Statistics: How Many People Use Twitter?," Backlinko, 2024. [Online]. Available: https://www.backlinko.com/twitter-users.