Article

Geospatial Mapping of Suicide-Related Tweets and Sentiments among Malaysians during the COVID-19 Pandemic

Noradila Rusli [1,2,*], Nor Zahida Nordin [2], Ak Mohd Rafiq Ak Matusin [1,2], Janatun Naim Yusof [3], Muhammad Solehin Fitry Rosley [4], Gabriel Hoh Teck Ling [5], Muhammad Hakimi Mohd Hussain [6] and Siti Zalina Abu Bakar [6]

1. Centre for Innovative Planning and Development (CIPD), Faculty of Built Environment and Surveying, Universiti Teknologi Malaysia, Johor Bahru 81310, Johor, Malaysia
2. Geospatial Research in Spatial Planning (GRiSP), Urban and Regional Planning Programme, Faculty of Built Environment and Surveying, Universiti Teknologi Malaysia, Johor Bahru 81310, Johor, Malaysia
3. Landscape Programme, Faculty of Built Environment and Surveying, Universiti Teknologi Malaysia, Johor Bahru 81310, Johor, Malaysia
4. Centre for the Study of Built Environment in the Malay World, Faculty of Built Environment and Surveying, Universiti Teknologi Malaysia, Johor Bahru 81310, Johor, Malaysia
5. Urban and Regional Planning Programme, Faculty of Built Environment and Surveying, Universiti Teknologi Malaysia, Johor Bahru 81310, Johor, Malaysia
6. Malaysia Urban Observatory Unit, Federal Department of Town and Country Planning (PLANMalaysia), Ministry of Housing and Local Governance, Federal Government Administrative Centre, Putrajaya 62675, Malaysia
* Correspondence: noradila@utm.my

Abstract: The government enacted the Movement Control Order (MCO) to curb the spread of the COVID-19 pandemic in Malaysia, restricting movement and shutting down several commercial enterprises around the nation. The crisis, which lasted over two years and featured a few MCOs, had an impact on Malaysians' mental health. This study aimed to understand the context of using the word "suicide" on Twitter among Malaysians during the pandemic. "Suicide" is a keyword searched for on Twitter when mining data with the NCapture plugin. Using NVivo 12 software, we used the content analysis approach to detect the theme of tweets discussed by tweeps. The tweet content was then analyzed using VADER sentiment analysis to determine if it was positive, negative, or neutral. We conducted a spatial pattern distribution of tweets, revealing high numbers from Kuala Lumpur, Klang, Subang Jaya, Kangar, Alor Setar, Chukai, Kuantan, Johor Bharu, and Kota Kinabalu. Our analysis of tweet content related to the word "suicide" revealed three (3) main themes: (i) criticism of the government of that day (CGD) (N = 218, 55.68%), (ii) awareness related to suicide (AS) (N = 162, 41.44%), and (iii) suicidal feeling or experience (SFE) (N = 12, 2.88%). The word "suicide" conveyed both negative and positive sentiments. Negative tweets expressed frustration and disappointment with the government's response to the pandemic and its economic impact. In contrast, positive tweets spread hope, encouragement, and support for mental health and relationship building. This study highlights the potential of social-media big data to understand the users' virtual behavior in an unprecedented pandemic situation and the importance of considering cultural differences and nuances in sentiment analysis. The spatial pattern information was useful in identifying areas that may require additional resources or interventions to address suicide risk. This study underscores the importance of timely and cost-effective social media data analysis for valuable insights into public opinion and attitudes toward specific topics.

Keywords: mapping; geospatial; GIS; suicide; Twitter; social media; data mining; sentiment analysis; spatial pattern

1. Introduction

Social distancing, masked faces, school closures, working from home, and striking images of empty city streets, parks, beaches, plazas, and walkways became a norm worldwide during the COVID-19 pandemic [1,2]. The pandemic impacted daily life unprecedentedly, resulting in impaired physical movement, restricted social connection, and mental struggles [3]. In 2020, the world's life expectancy dropped by 2.9 years, and the death toll increased in an overwhelming way [4]. For instance, based on the fatality rate, Europe was badly affected by COVID-19 compared to other nations in the minority world/global north [4,5]. Previous studies concerned the fatality rate because deaths were concentrated over months rather than several decades [6]. Unlike Europe, the fatality rate in Asia was not as high. However, the World Economic Forum predicted that most Asian countries would take longer than Europe to recover [7] in the financial, health, social welfare, and wealth-being sectors. While the short-term effect of COVID-19 focused on the fatality rate, it does not mean that the long-term effect, which includes the social well-being related to mental health that reflects one's emotions, psychology, and social well-being, should be neglected. The state of mental health reveals how an individual thinks, feels, and acts, carrying long-term pandemic effects [7,8]. Mental health strongly impacts how an individual interacts with others, deals with problems, and makes daily decisions about complicated circumstances [8,9].

According to research undertaken by the World Health Organization (WHO) in 2020, the pandemic caused a rise in mental health issues, such as depression and anxiety, known suicide risk factors. According to the study, the epidemic caused disruptions in daily life, financial hardship, and a lack of social support, which might increase suicide risk [10]. Suicidal thoughts increased among persons with pre-existing mental health disorders due to the pandemic. During the pandemic, those with depression, anxiety, and post-traumatic stress disorder (PTSD) had an increased incidence of suicidal ideation [11]. According to the WHO, suicide is "the act of intentionally causing one's own death." [12]. Meanwhile, ref. [13] defines suicide as a self-injurious behavior and not a disease, which includes suicide and attempts (inferred or actual intent to die). Ref. [14] sees suicide as the act of killing oneself. It is a major public health concern and a leading cause of death globally. It is a complex issue, and the reasons for suicide may vary from person to person. Factors contributing to suicide include mental health conditions, such as depression, anxiety, and substance abuse, as well as life events, such as the loss of a loved one, financial difficulties, and relationship problems.

A few factors may lead to this state. In the context of psychiatry, the transition from suicidal ideation to a suicide attempt or action is facilitated by the co-occurring psychiatric conditions that increase distress (such as panic disorder or PTSD) or decrease restraint (such as substance abuse or cluster B personality disorders) [13]. Simultaneously, additional factors involved in this transition include the capability for suicide, exposure to suicide, mental imagery, and access to the means of suicide [15–17]. Moreover, refs. [18,19] indicate the factor by linking behavioral traits with long-term suicidal behavior. The exposure or past experience of suicidal behavior could trigger the victim's intention and suicide attempt in the future if it is not treated properly. Anxiety might also contribute to suicidal behavior by interacting with other characteristics of individuals with an increased vulnerability to suicide [13].

In the medical context, transmitting impulsive aggressive traits partially explains familial aggregation of suicidal behavior [20]. Studies on suicidal behavior heritability found evidence that suicide and suicide attempts are transmitted independently of psychopathologies [21]. Furthermore, the offspring of individuals who have attempted suicide have a five-fold risk of attempting suicide than the general population [22]. The risk of suicide is double among individuals who lost a parent suddenly due to other causes [23]. However, no single gene or group of genes has been identified as responsible for suicidal ideation, suicide attempt, or suicide across multiple studies [24].

There is a scientific relationship between mental health and suicide. Studies have shown that individuals with mental health conditions, such as depression, bipolar disorder, and schizophrenia, are at a higher risk of suicide. Individuals with depression have a 15-fold increased risk of suicide compared to the general population [25]. The study also found that individuals with bipolar disorder and schizophrenia have a 17-fold and 6-fold increased risk of suicide, respectively. Another study found that individuals with major depressive disorder (MDD) have a 20-fold increased risk of suicide compared to the general population [26]. It also found that individuals with MDD, who also have a comorbid anxiety disorder, have an even higher risk of suicide. A published systematic review found that individuals with PTSD have a 2- to 4-fold increased suicide risk compared to the general population [27]. It is important to note that suicide is not a normal response to stress, and mental health conditions should be treated as soon as possible.

In Asia, especially in southeast Asian countries such as Malaysia, the number of reported COVID-19 cases significantly parallels the number of reported suicides. By 2020, there was an increase of almost 44%, equal to 631 suicide cases recorded compared to only 356 from 2014 to 2019 [28]. Moreover, there is a significant association between reported news on suicide, mental health, and depression and total COVID-19 cases in Malaysia. For example, in March 2020, when 3236 new COVID-19 cases were reported, there was zero news on suicide, mental health, and depression. However, in July 2021, when 361,293 new COVID-19 cases were reported, the news on suicide, mental health, and depression rapidly increased to 129. This shows a significant correlation between new COVID-19 cases and the number of news stories on suicide, mental health, and depression (Figure 1).
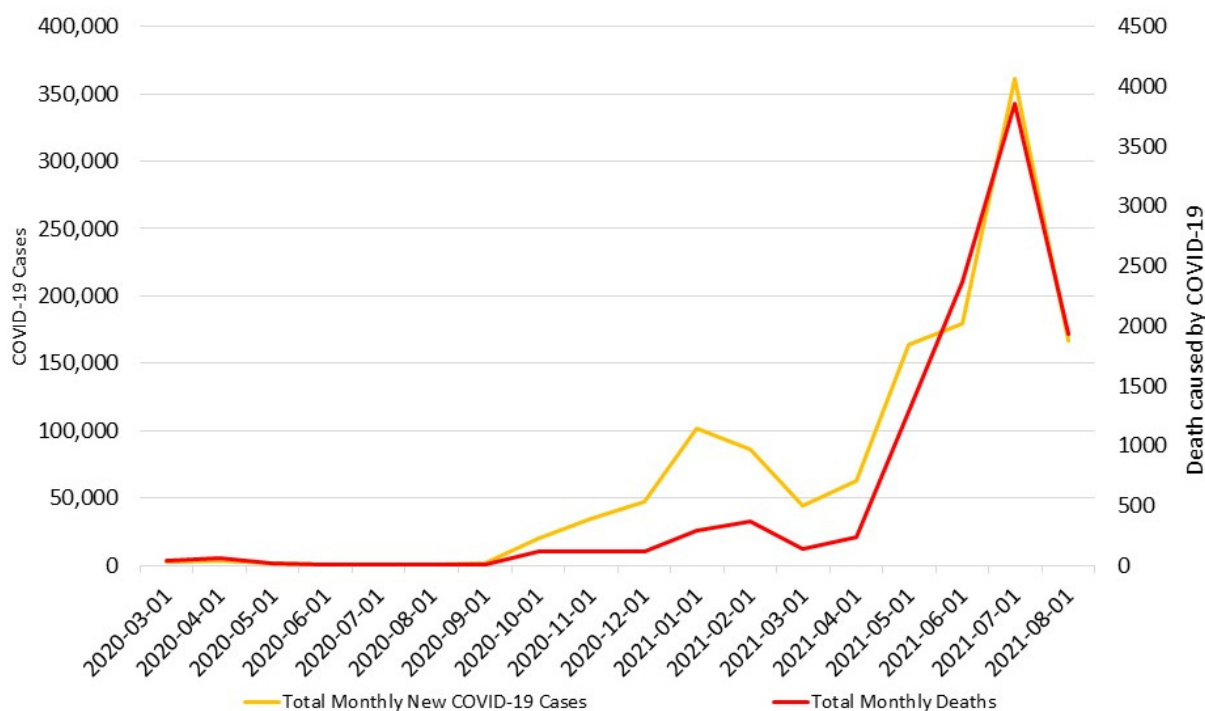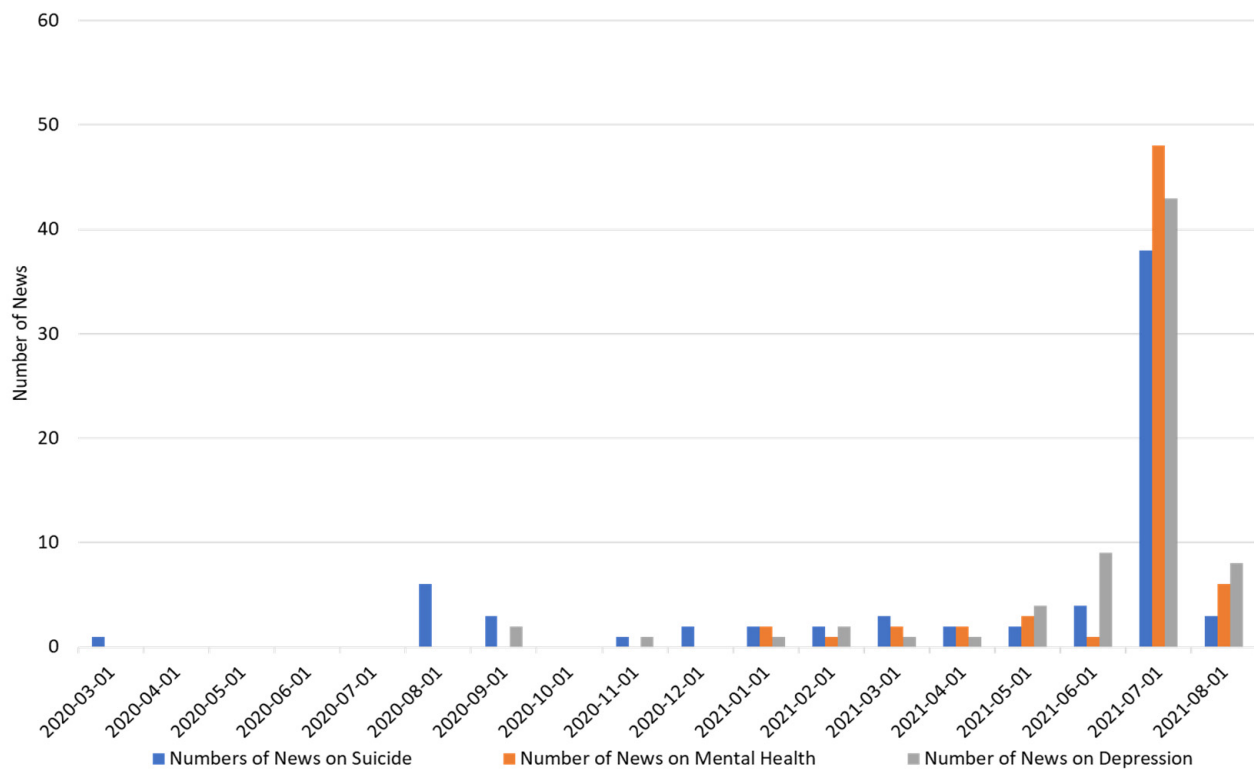

Figure 1. *Cont.*

Figure 1. Statistics of the comparison between COVID-19 new cases and deaths as well as the number of news stories related to mental health, suicide, and depression. Note: Data on the number of news stories related to mental health, suicide, and depression were collected by the author using a data mining technique. Statistics on COVID-19 cases and deaths were obtained from the Ministry of Health [29].

Most research on mental health issues has been conducted predominantly in Europe and the North American continent. There are few representations from Asian countries, especially Malaysia, which is in southeast Asia. This suggests a lacuna and more research in Asian countries is desperately needed for knowledge translation. It is imperative to understand how far COVID-19 has impacted the world's inhabitants, specifically in Malaysia, by mining and interpreting big data from social media sources. Generally, big data comprises three dimensions: high volume, high level of variety, and high velocity [30,31]. Text, images, audio, and video are examples of unstructured data hoarded from internal sources (e.g., sensor data) and external sources (e.g., social media) [32,33]. Malaysia ranks among the top three for the highest time spent on social channels among Asian countries. Malaysians use social media for 3 h and 1 min per day and rank 13th among social media users [34,35], which means we can extract unstructured data from external sources of Malaysian social media users to understand their virtual behavior. It also suggests that social media, such as YouTube, Instagram, Twitter, Facebook, TikTok, and WhatsApp, are platforms where most virtual communities in Malaysia share their interests, concerns, and views. These were delivered through liberation in expressing feelings and thoughts by posting their status on their page, tweeting others' statuses, uploading videos and photos, commenting, and sharing the content [36]. These virtual social phenomena created virtual behaviors [37] and, eventually, turned social media platforms into a new virtual social milieu that changed Malaysians' daily social landscape.

However, how have these social phenomena created a new mode of communication related to mental health, depression, and suicidal sentiments among Malaysian during the spark of the COVID-19 pandemic? The relationship can be understood from two standpoints: the MCO chronology in Malaysia and the active social media users as a percentage of the total population in Malaysia from 2016 to 2021. The government declared the MCO from 18 March 2020 to 31 March 2020 to contain the COVID-19 outbreak, effectively slowing down transmission rates. The government proclaimed it under the Prevention and Control of Infectious Diseases Act of 1988 (Act 342) and the Police Act of 1967. In total, MCOs lasted for seven weeks, consisting of four phases from MCO 1 to MCO 4 (from 18 March 2020 to 3 May 2020), followed by a six-week Conditional MCO (CMCO) (from 4 May 2020 to 9 June 2020) and subsequently Recovery MCO (RMCO) (from 9 June 2020 to 31 August 2020) [38]. The government implemented various non-pharmacological interventions during MCO 1, which continued during MCO 2, and gradually relaxed rules in subsequent MCO phases. During MCO 3, there were relaxations in the MCO directives, where some sectors were permitted to operate, e.g., certain construction projects and the automotive industry [38]. However, permissions were subject to approval from the Ministry of International Trade and Industry (MITI) to reduce the number of employees at the workplace. The further expansion allowed almost all economic sectors to function fully during MCO 4. During these periods, most Malaysians spent long hours alone or stayed with their family at home, limiting their face-to-face social engagement with friends, family members, and colleagues.

As of January 2021, about 86% of the Malaysian population were active social media users, which increased by 24% from 2016 to 2021 [35]. In 2016, social media users accounted for about 62% of the Malaysian population. Facebook was the leading social media platform. When comparing the increase in social media penetration among the Malaysian population, very little change was observed, with an average gain of 3% from 2017 to 2019. However, during the MCO from 2020 to 2021, the values rose to 5% for two years, and it turned out that 86% of the population were active social media users [34,35]. This denotes that when most Malaysians had to stay and work from home, they increased their screen time during the MCO. They chose social media platforms to replace or replicate their previous face-to-face social engagement. For example, it was not limited to existing friends, family members, and colleagues. They connected with anonymous people they considered new friends by sending friend requests or following their pages [39]. Sometimes their profile pictures, bios, and comments became valid, knowledgeable, and convincing when discussing certain issues [40], which led to social engagements and companionships that were no longer limited to day-to-day physical face-to-face encounters.

Therefore, this study aims to gain a deeper understanding of the context in which Malaysians used the word "suicide" during the COVID-19 pandemic by using big data mining from a social media platform, i.e., Twitter. This paper begins by addressing the disturbing subject of high suicide incidence in Malaysia during the COVID-19 outbreak. As COVID-19 mortality increased, there was a linear increase in suicide cases. An analysis of Twitter content is necessary to understand better how Malaysians used the term "suicide" during this period. The study looked into regional patterns and trends in suicide on Twitter during the pandemic to discover potential risk factors and insights we might use for public health interventions to lower suicide rates in Malaysia.

2. Materials and Methods

We chose both Peninsular Malaysia and Borneo as the study area. Malaysia has a population of over 32.7 million. Not everyone uses social media due to age restrictions and advanced technologies. Normally, only those over 18 use it. We paired social media and NVivo 12 [41,42] to investigate suicide sentiment in Malaysia. Social media is a source of big data, and NVivo is a tool for mining it. Tweeps generate a large amount of data. The NVivo 12 plugin also extracts the users' location. NCaptures is a free web browser extension developed by QSR that enables gathering material from the web for import into NVivo 12. NCapture can collect various content such as articles, blog posts, Facebook,

Twitter, and YouTube. Through NCapture, NVivo12 can capture, extract, and mine the spatial location of the tweeps (latitude, longitude) [43,44].

Although the time frame for data mining settings extends from March 2020 to August 2021, data collection was only available from 3 to 6 August 2021. We extracted 18,000 tweets worldwide containing the keyword and hashtag "bunuh diri" (in English: suicide). We used the Malay language keyword to focus the data mining on Malay language users and their locations because, given the large number of English-speaking users, the data would undoubtedly be higher if we used English words. It is crucial to remember that NCapture does not gather every tweet. Based on the number of Tweets and Twitter traffic at the collection time, Twitter's API determines the random subset of Tweets that were made available. Additionally, the period during which we could gather tweets and the maximum number of tweets we could collect at a time (max = 18,000 per 15 min) were constrained by Twitter's API (only tweets from the past seven days) [45]. The period covered by the data mining on Twitter is consistent with previous research on Twitter analysis [43,44,46].

Data Preprocessing and Analysis

We did not consider tweets from countries other than Malaysia and determined location based on geographical information and language. Because the keyword "bunuh diri" is also used in Indonesian, a significant amount of collected data were tweets that originated in Indonesia. By applying a filtering procedure based on geographical location and tweet content relevance, we included only 392 tweets with latitude and longitude information from Malaysia in the analysis. The limited number of tweets is consistent with previous studies [43,44,46], which also restricted their analysis to tweets specifically related to the context of their study.

Within the Twitter application programming interface, it gives the following details on how latitude and longitude coordinates are assigned to Tweets: "If the user provides an accurate location, the coordinates object will return a (longitude, latitude) array containing the geographical coordinates, and a Twitter place that corresponds to that location will be assigned" [45]. The precision of these locations is unknown. Technically, Twitter may receive information about the user's location, such as the IP address, precise location information from the Global Positioning System (GPS), or information about wireless networks or cell towers near the user's mobile device [47]. We also used users' locations to determine the country setting for their accounts based on the countries where Twitter is available. If Twitter cannot resolve a user's location, it defaults to the United States. NCapture captures the location specified by the user based on the public profile, where the user can decide to share or make their location private. A user's country setting is private information that Twitter uses to modify the "availability" content for users. Interestingly, Twitter automatically marks Sarawak as being in the middle of Malaysia if the location specified by the user in their public profile is Malaysia. However, if the user sets the exact location, for example, Kuala Lumpur, the location will fall on the precise coordinates of Kuala Lumpur.

Next, the preprocessing step removes hashtags, usernames, URLs, [48,49] stop words, lower-case letters, title postfixes, and punctuation from the text (Twitter). Using the Google Collab platform, we created a text filtering Python program to incorporate NLTK (Natural Language Toolkit), a platform for building Python programs to work with human language data [50]. Then, the list of tweets originally written in Malay was translated into English while ensuring that each statement's context remained unchanged and conveyed the same meaning as it did in Malay. Per Twitter data analysis best practices and media user privacy, we eliminated usernames, biographical details, and other potentially identifying information. The data collected is known as secondary data and is called social media monitoring. We chose this social media monitoring technique because it fits the scope of our study and aims. Social media postings are from users sharing and expressing their sincere thoughts, opinions, and views on a topic. For this technique, the sample size does

not matter because it depends on the number of datasets collected by web crawlers and the keywords used in this study.

This study also involved sentiment analysis ("opinion mining"), which used natural language processing, text analysis, computational linguistics, and biometrics to systematically identify, extract, quantify, and study affective states and subjective information. Valence Aware Dictionary and Sentiment Reasoner (VADER) is a rule-based sentiment analysis tool and a lexicon used to express sentiments on social media. It is an open-source tool (https://github.com/cjhutto/vaderSentiment, accessed on 15 December 2022) developed by [51]. It has a high classification success [52] when analyzing emotions of social media text, which is useful for studying sentiment derived from Twitter data [53]. We used the VADER Sentiment Analyzer to classify the preprocessed text as positive, negative, or neutral. The typical threshold values used are:

Positive: compound ≥ 0.05;

Neutral: (compound > -0.05) and (compound < 0.05)

Negative: compound ≤ -0.05

We calculated the compound score by adding the valence ratings of each word in the lexicon, adjusting them according to the rules, and then normalizing them to fall between -1 (the most extreme negative) and $+1$ (the most extreme positive). This is the most relevant metric if one unidimensional assessment of a sentence's sentiment is required. We utilized VADER below normalization [51].

$$x = \frac{x}{\sqrt{x^2 + \alpha}} \tag{1}$$

where x = sum of valence scores of constituent words, and α = Normalization constant (default value is 15).

Then, we validated the results using evaluation metrics: precision, recall, and F score. The equation for calculating those metrics is as follows [54]:

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \tag{2}$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \tag{3}$$

$$\text{F score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

The metrics were computed using a total of 100 messages. While there is no one-size-fits-all technique for establishing sample size in sentiment analysis studies, a basic rule of thumb that is frequently referenced in the literature is to use at least 100 messages [55,56]. Human raters were used to annotate the messages by establishing some criteria to consider human raters, which is shown in Table 1. Two human raters were used to validate the sentiment analysis results.

Table 1. The criteria of human raters.

Criteria	Description
Language proficiency	Understand the language used in the messages to be annotated thoroughly. They should be able to interpret nuances in meaning and tone and be conversant in the language. The raters were given the 100 messages in both languages; English (with sentiment analysis value) and Malay (to interpret nuances in meaning and tone).
Consistency and reliability	The raters' annotations should be consistent, and they should be able to achieve a high level of inter-rater reliability. Inter-annotator measurement metric [55,57] was calculated using Cohen's kappa, and the value is 0.82 for human rater number 1 and 0.8 for human rater number 2, which consider an almost perfect agreement.
Objectivity and impartiality	The raters' annotations must be neutral and fair, with no biases or prejudices that might influence their conclusions.

During the validation process, disagreements between human raters were solved by discussions among raters and an attempt to reach a consensus. However, for unsolved disagreements, a third party, which in this case is the leader of this research, will make a final decision on the sentiment category by relying on the annotators' judgment.

This study also incorporated content analysis techniques to analyze text data through the systematic classification process of coding and identification of themes and patterns [58]. We present the qualitative analysis results in categories, which allow interpretation of the text [59]. This analysis aims to identify the theme of tweet discussion by the users. We performed these analyses using NVivo 12. Before the coding process began, we examined the themes together to comprehend the data pattern. Then, we reviewed the theme of tweet discussions to identify meaningful subjects that corresponded to the purpose of this research.

We used content analysis to support sentiment analysis. A major drawback is that sentiment analysis is limited to native English speakers. Therefore, it does not distinguish between sentiments expressed toward different targets when non-native speakers of other languages express themselves. Moreover, this research also identified the differences in how people talk about themselves in various domains related to cultural relativism. For example, Malaysians synonymously engage in sarcastic sentiments by expressing their positive feelings using the text's constructive or intensified positive words. This sarcasm is an indirect expression, but Malaysians fully understood that the content was a reprimand. Using sentiment analysis without fully comprehending the context and situation of the tweets would most likely classify the sentence as expressing positive sentiments because of the presence of words such as "meningkat" and "sokongan", which directly translate as "increase" and "support". The truth is that tweeps express their negative sentiments using sarcasm. Hence, using this constraint, we computed the word cloud using NVIVO 12 Plus. This type of content analysis guides the researcher's perspective to understand the kinds of situations in which words are used most often. This cleared up the researcher's confusion about why people use positive or intensified positive words to express their negative feelings in tweets. In other words, the highest ranking words generated in the word cloud were double-checked and coded according to negative or positive themes. Figure 2 shows the overall methodology of this study.
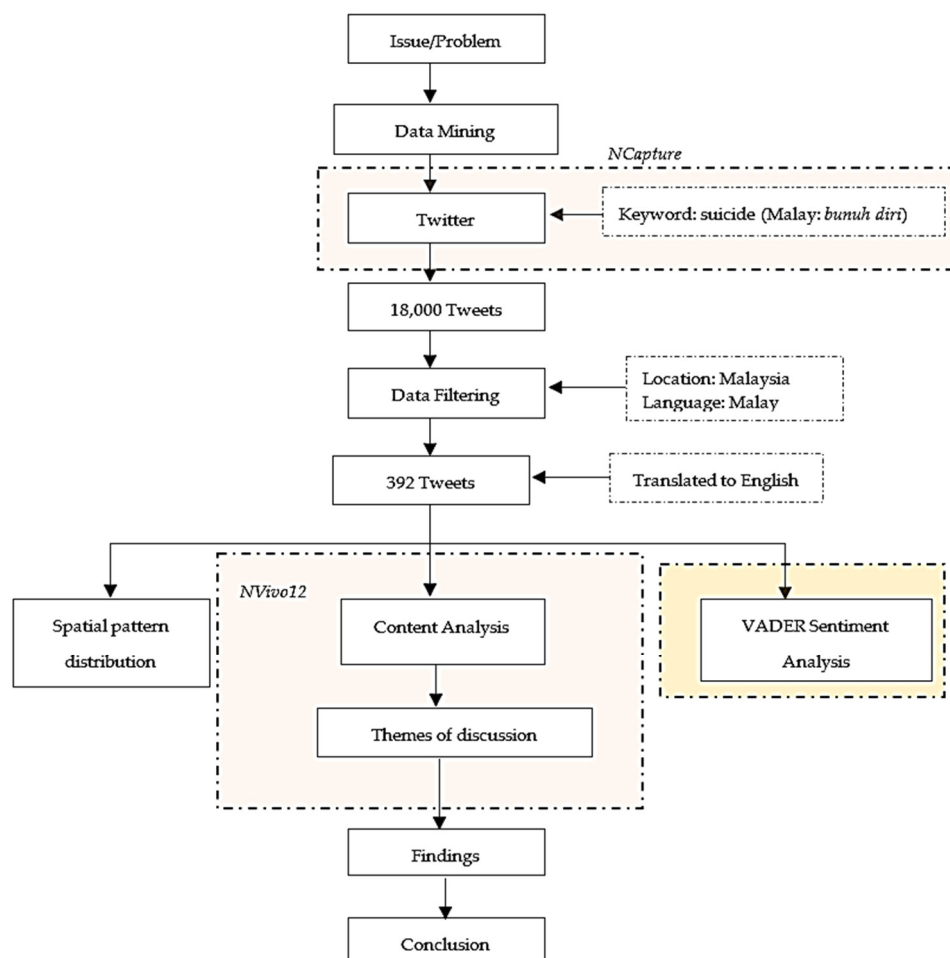
Figure 2. Overall research methodology.

3. Results and Discussion

3.1. Thematic Content

We analyzed a total of 392 tweets, divided into three categories: "Critics of the Government of the Day (CGD)" (N = 218, 55.68%), "Awareness of Suicide (AS)" (N = 162, 41.44%), and "Suicidal Feeling or Experience (SFE)" (N = 12, 2.88%). The CGD theme was tweeted about current standard operating procedures (SOP) implemented by the government, MCO, and political instability. Meanwhile, the AS theme became the topic with content such as suicide statistics, suicide factors, and debate on the low faith of suicidal people. In the following, we interpreted "SFE" as (i) Twitter users feeling suicidal while tweeting, (ii) Twitter users' suicidal experiences in the past, and (iii) experiences by close people about suicidal feelings and suicidal incidents. Also, we are not restricted to a single theme within each tweet. It is also possible to have two or three themes running simultaneously. Table 2 shows examples of tweets.

3.1.1. Critics of the Government of That Day (CGD)

Normally, suicide ideation and thought entail actively preparing for suicide, although suicidal thoughts do not always mean a person intends to take their own life. Nevertheless, when Malaysian users tweeted this word publicly as a hashtag, it revealed new directions and convictions. From the Malaysian perspective, using the word "suicide" was interpreted and depicted in different perspectives and situations, depending on their current and daily circumstances. One of them was a critic of the government. Most users expressed frustration and disappointment with how the government handled the nation's economic recession caused by the pandemic, which affected them. Political instability led to economic

instability, resulting in loss of employment and income, fear of COVID-19 infection and death, and students' frustration with remote learning. This was reflected in users' tweets of the keyword "suicide", which they posted amid the rising number of COVID-19 cases (more than 20,000 cases) and more than 350 deaths [29].

Table 2. Theme and examples of tweets.

Theme	Description	Example
Critics of the government of that day (CGD) Total: 55.68%	Criticism of the government's strategy in dealing with the spread of COVID-19 in Malaysia.	All the strategies have failed. Has SOP yet been violated? Specialist advice has been ignored. Ignored the request to shut down factories. People are currently jobless and homeless. Business owners are falling into bankruptcy. Family problem.
	Sarcastically criticize the Malaysian government's response to the spread of COVID-19.	Thank you, KKM, for giving this info. After all, the people didn't even know thinking about the causes of these clusters and some even committed suicide. Finally, KKM shared that the cause is from factories. Thank you, KKM, for the info.
	Condemning those who continue to support the government	I understand if there are still T20 classes supporting the government. Maybe they have their own benefits that need to be taken care of. But why are the B40 classes still supporting them? Don't you guys see the destruction happening in this country alone?
	Political comments	People die almost 200+ every day. Almost 1 million people lost their loss of income. 1000 people committed suicide. Our ministers were clapping their hands because the PM had secured his position. This is so cruel.
Awareness related to Suicide (AS) Total: 41.44%	Concern for the country's situation	Success Story -COVID 16,000+ per day -Death due to COVID 200+ per day -1000+ suicide -The nation's poverty rate increased by double. -More than 500,000 of M40 families currently hold the status of B40. -1/4 graduates unemployed today!
	Child marriage and polygamy	1. Child marriage with forcing the children. The kid had been suppressed before committing suicide. Then is it the kid's fault? 2. A husband can marry polygamy without the wife's permission, so it needs to be approved? Just throw them away if it's not good.
	Religious	Don't do something that will send us to Allah's hell. A video of the suicide of a mother and her two children really touches Sufi's feelings and turns into an inquiry of Sufi: Where is our *imaan* until we are willing to put our religion behind?
		In Friday's sermon at the mosque, I can hear it from my house. He says that those who commit suicide by drinking poison will be punished and tortured by drinking poison forever in the afterlife.
		Suicide cases are increasing. Please do know about it. Don't judge. Even God is the one who is Loving and Forgiving.

Table 2. *Cont.*

Theme	Description	Example
Suicidal Feeling or Experience (SFE) Total: 2.88%	Personal experience	Once in this situation, I was determined to commit suicide. Medicine overdose. Could not accept the truth.
	Feeling depressed	I once had that experience, but not to the level of committing suicide. I just feel depressed and so lonely, like there is no one to share problems with. But suddenly, I met the boys and followed them. They're very helpful.
	Others' experience	My supplier has been talking about committing suicide because of depression with company problems that they had to face. Their sales are 70% affected due to the pandemic. Even though they are selling essential items, they are still affected.
		My uncle sent a photo of one Indian uncle committing suicide at Bukit Belimbing, which is very near to my house. O' Allah, I'm so sorry for him. He is alone and unable to cope with the stress of life.

Disclaimer: This study reflects the opinions and sentiments expressed by the participants towards government bodies and religion. The views presented in this study are not intended to cause harm or offense to any individual, group, organization, or religion. The study is purely for academic purposes and does not represent the views of the researchers or the institutions they are affiliated with.

Several incidents led to this discontentment, e.g., political stability affects economic stability. The pandemic and Malaysia's political unrest happened at the same time. The political chaos in Malaysia, which led to Mahathir Mohammad's resignation as prime minister in February 2020, included the "Sheraton move." Muhyiddin Yassin succeeded him in establishing a new government on 1 March 2020. Malaysia's political unrest continued in 2020 and 2021. As a result, Muhyiddin Yassin's administration fell in August 2021, and a new government was installed under the leadership of Ismail Sabri Yaakob. In Malaysia, the pandemic had catastrophic economic and societal impacts. Malaysia's GDP decreased as a result of the decline in the value of its currency. As a result, the government's first reaction to the pandemic was delayed. Subsequent waves of COVID-19 infections and emergency declarations worsened the nation's ongoing political instability during 2020 and 2021.

People were not allowed to leave their homes to stop the spread of COVID-19, but the industrial sector was still open and was a major factor in the rise of COVID-19 cases. The government's strategy seems to favor the wealthiest, who were from the social class T20 while disregarding the plight of the less fortunate, who were from the social class B40. Furthermore, the pandemic affected multiple levels of society due to national political instability. The growing number of cases sparked public outrage at how the government handled the situation, dismissing dissent among Malaysians. As a result, tweeps used Twitter to express their anger. Therefore, we identified that the keyword "suicide" presented emotional frustration that displayed discontentment and flamed public anger, which means that suicide is a poly semiotic word used to elicit the user's emotions. This is in line with the finding by [60], who mentions that, compared to those who submitted messages with more positive posts, those who wrote tweets with negative posts tended to reach a larger audience.

Paradoxically, the keyword "suicide" also conveyed a positive message, and there were tweets that encouraged positivity and an understanding of mental health to support those who needed it. Despite the circumstances under which most people were forced to express their displeasure, the keyword also delivered messages in the form of prayers, motivational quotes, discussions, and information sharing. This led to relationship building and maintaining connections with others that could help spread happiness and lessen feelings of loneliness, meaning that some tweeps were compassionate and empathic toward others by giving assistance, disseminating information, and highlighting others' achievements.

This reveals that the keyword "suicide" conveyed new convictions of positivity by complimenting people, highlighting their efforts, and exhibiting compassion through Twitter. The keyword "suicide" did not solely reflect negative emotions but also positive ones. The word "suicide" should be viewed from different angles in its local context of origins, countries, and social cultures in a particular milieu where people deal with daily difficulties.

3.1.2. Awareness Related to Suicide (AS)

Tweets by Malaysians also expressed concern about suicide cases that they believed correlated with the high number of COVID-19 cases. The tweets included statistics on COVID-19 cases and deaths and the high number of suicide cases. Additionally, there were concerns about the socioeconomic situation in Malaysia because people were becoming poorer due to the economic problems caused by the pandemic. According to the Department of Statistics, the unemployment rate in Malaysia increased in recent years: 3.3% in 2019 [61], 4.5% in 2020 [62] and 4.6% in 2021 [63]. The increase in the unemployment rate could contribute to the economic problems, and people were concerned that more middle-class families (M40) would become lower-class (B40) if the government did not take action to mitigate the economic problems.

Additionally, there were tweets offering religious advice on the repercussions of suicide. People should consider their *iman* (faith) and religion before taking their own lives. Many religions view suicide as a grave sin and a major offense. For instance, Islam considers suicide a grave sin, and those who do so will face punishment in the hereafter. This point of view is based on the notion that God gave us life as a gift and that people should not take their own lives.

Malaysians also use Twitter to share hope and inspiration and let people know that support is always available. These messages might emphasize the importance of religion and faith in trying times and offer advice on dealing with hopelessness and despair. It is critical to remember that suicide is a complex issue with no single cause or justification for why someone might think about committing it. Risk factors include mental health challenges, money and relationship troubles, and a history of abuse or trauma. Furthermore, many reasons may play a role in a person's decision to take their own life, and it is important to realize that nobody is ever truly alone.

3.1.3. Suicidal Feeling or Experience (SFE)

In the Malaysian context, the keyword "suicide" also carries implicit and explicit expressions of suicidal feelings, past experiences with suicide, and observations of close ones experiencing suicidal thoughts or incidents. The word cloud (Figure 3) shows that the keyword "suicide" was always partnered with "covid", indicating that the pandemic triggered a new dimension of emotional states of helplessness, demotivation, less physical interaction, and a low willingness to live a daily routine. However, it is too early to conclude that someone who used the words "suicide" and "covid" had the ideation of suicide, which indicates that by understanding the keywords tweeps used, Twitter had become an informal resource. This concurs with [63], which proves at-risk people use modern technology, such as forums or microblogs, as an informal resource to communicate their greatest difficulties without speaking to anybody in person. Due to the social stigma associated with mental diseases, at-risk individuals may forego receiving professional help. Therefore, understanding the keywords Malaysian tweeps used could effectively spot virtual behavioral and emotional state changes in potentially at-risk people at the early stages.
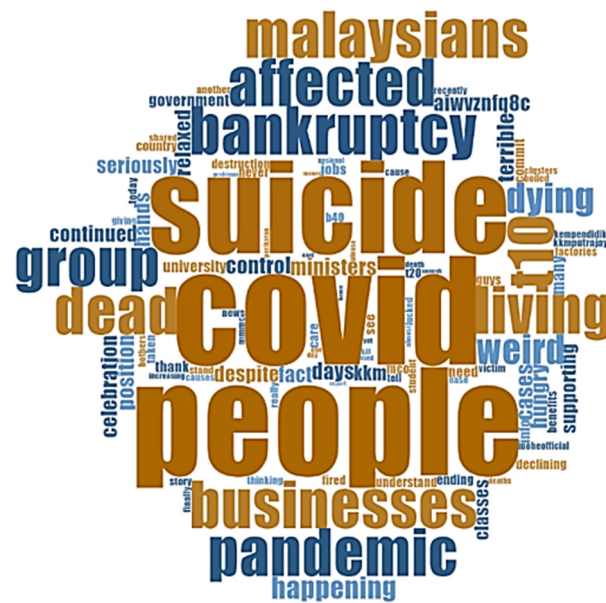
Figure 3. The word cloud in Twitter after content analysis.

Each tweet shows the user's feelings, disappointment, anger, and anxiety regarding the pandemic situation. For example, "covid" was mainly used to express the issue related to the increasing number of infected cases and deaths. This could be why the word "covid" was used more in tweets posted to express and discuss the current situation. "Suicide" could express the reported suicide cases and suicide attempts co-occurring with COVID-19 because the MCO and lockdown period, which lasted almost two years, had various effects, especially on social life. Most people felt helpless and unmotivated, had less physical interaction, and had a lower willingness to live, which may be the reason for an increasing number of reported suicide attempts. Referring to the Penang Bridge suicide attempt, there were many reported cases of suicide and attempted suicide during MCOs. Thus, these were issues that raised concern and drew the attention of our society and the authorities, leading to a hotline focused on assisting suicidal people. Because we focused only on the characteristics of the keyword used, there is a strong need for a novel approach, especially in quantifying the risk level for an individual. For instance, based on a user's behavior intensity on Twitter, which includes speech, daily activities, and the size of a social network, we may derive a Suicide Prevention Assistant (SPA) text score. SPA is required to interpret the abundance of information on people's emotions and actions so that it can be utilized effectively to spot behavioral changes in at-risk individuals early on and possibly even save lives.

3.2. Tweet by Location

We classified tweeps' location into five zones: west, east, north, south, and Borneo (Sabah and Sarawak) (see Figure 4). We used Geographic Information System (GIS) tools, such as ArcGIS, to show the locations of the tweets in a geographical context. We used GIS tools to plot each tweet's latitude and longitude information and show them as points on a map. We used cluster analysis to discover data patterns. Cluster analysis is a statistical approach for grouping similar data points together, which can assist in finding trends and patterns in huge datasets. The cluster analysis model employed in this investigation was [64]:

$$\text{Euclidean Distance} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \qquad (5)$$

where x_1 and y_1 indicate one tweet's latitude and longitude, and x_2 and y_2 represent another tweet's latitude and longitude.
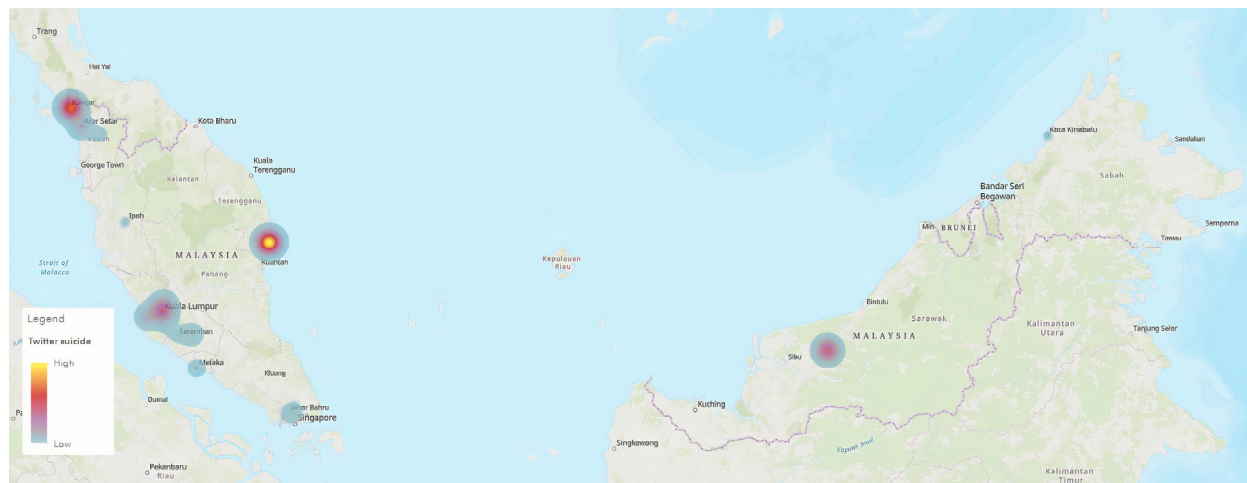
Figure 4. The heatmap of tweets' location related to suicide content.

The location information for the 392 tweets is shown as clusters on the map using this approach, with different tones signifying high and low values. This enables a visual representation of the tweets' geographic distribution and highlight places with higher and lower tweet activity.

According to the breakdown, Borneo, which consists of Sabah and Sarawak, had the highest tweet rate of 35.6%. The western zone was second at 33.6%, the east at 16.1%, the north at 13.6%, and the south at 1.3%.

We found that several important cities in Malaysia contributed a high volume of tweets: Kangar and Alor Setar in the northern zone; Kuala Lumpur, Klang, and Subang Jaya in the west; Chukai and Kuantan in the east; Johor Bharu in the south; and Kota Kinabalu in Borneo. More affordable internet packages, faster connections, smartphone access, and the time urbanites spend on the Internet allowed social platforms to penetrate their daily lives. Thus, social media platforms became necessary and accessible for expressing emotions and feelings. Moreover, this could trigger a sense of togetherness between tweeps when they shared and retrieved information related to their interests. MCOs affected many people in these cities economically. Unlike other parts of Malaysia, these cities rely heavily on and serve as centers for hospitality, urban tourism, arts and culture, and manufacturing industries. These attracted people to migrate and live in these urban settings. MCOs halted almost all of these activities, leaving many urban residents unemployed. They posted about their anger, frustration, and hopelessness over this unprecedented economic regression on their social platforms, expecting emotional relief and support from other users.

The lack of communal engagement, particularly in urban areas during MCOs, might contribute to the high number of tweets. The sense of belonging due to the emotional support of virtual communities should replace physical, face-to-face communal support. Tweeting or highlighting "suicide" in an individual's news feed is one way to draw public attention. Some of these posts were retweeted or shared on social media platforms to increase people's awareness or criticize the government's current mitigation plans. Figure 5 compares the number of tweets among cities in Malaysia according to the tweet content.

Kuala Lumpur had the highest population density, with 7188 people per square kilometer [65], putting the city at a very high risk for the spread of COVID-19. For four months, from 16 June 2020 to 16 October 2020, the statistics displayed a constant cumulative number of COVID-19 cases. However, the following month, on 16 November 2020, the number of COVID-19 cases began to rise, and between 16 June and 16 July 2021, the statistic peaked (Ministry of Health Malaysia, 2022). This situation triggered various reactions among Kuala Lumpur's citizens. Based on the tweet statistics (Figure 5), 50% of the tweets were CGD, 44% were AS, and the remaining 6% were SFE. Most tweets discussed the issue of the higher number of COVID-19 cases and deaths, the increase in

suicide, and government incompetency in managing the country's needs and stability during the pandemic.
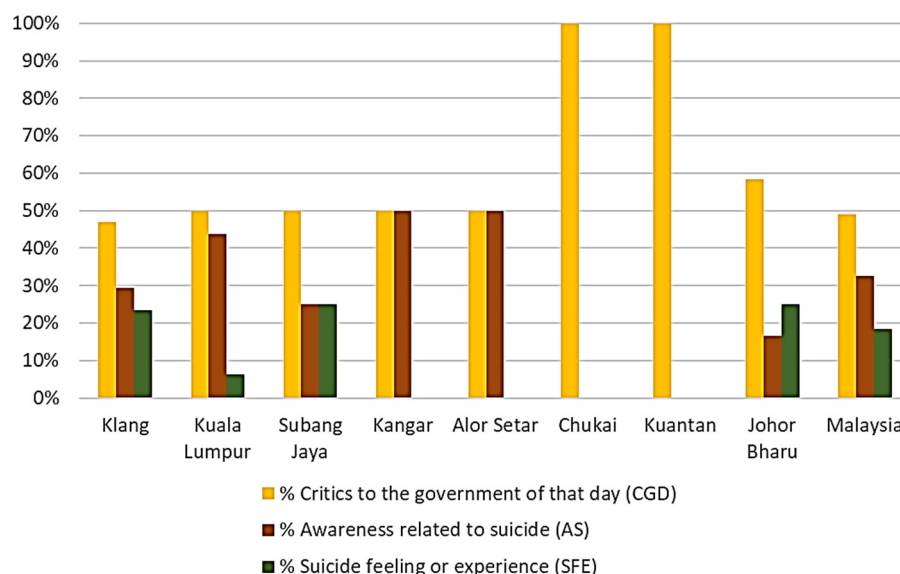

Figure 5. Percentage of tweets' content among cities in Malaysia.

Selangor is a highly populated state that also had many COVID-19 cases. Subang Jaya and Klang had more tweets, with the content in each tweet showing 50% were CGD, 25% were AS, and the remaining 25% were SFE. Those tweets mainly focused on criticizing the government, particularly on political incompetency and the loss of income many Malaysians faced. Klang recorded the highest number of COVID-19 cases (8000) in July 2021, which sparked tweets about the government. Tweeps posted about CGD against political instability using the hashtag "LAWAN" (fight). They expressed their disappointment and disagreement with the government's decision to continue the lockdown despite failing to significantly reduce the number of cases and help citizens: CGD (47%), AS (29%), and SFE. For instance, one tweet showed a suicide prevention hashtag with a suggestion to the government to help people in need. The tweet showed that people desperately needed support from the government but did not receive it. There were also instructions on how to share breaking news on suicide cases based on the severity of the condition during the lockdown. Users also shared their views and empathy towards suicide victims and hoped for the best for the people around them. Based on the postings, tweeps always channeled their negative and strong stance into showing their disappointment, protest, and disagreement with the government.

Meanwhile, Johor Bahru (located in the state of Johor, southern Malaysia) is one of Malaysia's most highly populated areas. From 16 December 2020 to 16 February 2021, Johor Bahru reported more than 1000 COVID-19 cases per day [29]. Similarly, due to many cases, citizens expressed their anger, disappointment, and protests toward the government through their tweets. 58% of the tweets were CGD, 17% were AS, and 25% were SFE.

Meanwhile, on 16 March 2021, Kangar (located in northern Malaysia) recorded a spike in the number of cases and a drastic drop a month later. This situation led to tweets about CGD and AS, with a fair percentage of tweets in both categories. The tweets were mainly users' sarcastic opinions about the sudden spread of the disease. It shows they were disappointed with how the government tried to hide the facts behind the spike. In Kedah, there were a drastic increase and decrease pattern in the number of cases, with an increase in July 2021, which created a negative reaction among citizens, mainly in the CGD. The critics expressed their disappointment with the government, especially with the effectiveness of the government's strategies in dealing with the pandemic.

In Pahang, the highest recorded cases were more than 600 on 16 July 2021. Even though it is in the first phase of the lockdown, reported cases stagnated from June 2020 to November 2020. However, the reported cases showed a roller-coaster pattern between 16 June and 16 July 2021. Due to the situation, the tweets posted in this area were CGD, which had a similar pattern of criticizing the government's capability in handling the pandemic uncertainties.

Terengganu similarly recorded high cases of COVID-19. The number of cases stagnated from April 2020 to January 2021. However, from the end of January 2021 to February 2021, the number of cases suddenly increased. Then the number rose and fell inconsistently between May 2021 and July 2021, which caused anxiety, anger, and disappointment among the citizens towards the government (CGD theme), reflected in 100% of tweets directed at the government.

3.3. Sentiment Analysis of Tweets

We validated the results from the VADER sentiment analysis using three (3) evaluation metrics: precision, recall, and F score. The evaluation yielded good findings with precision, recall, and F1 scores of 0.99, 0.94, and 0.96, respectively. These results demonstrate that the VADER sentiment analyzer executed the sentiment analysis task with great precision and dependability.

We can observe tweeps' sentiments from the time and date people post their tweets, which shows that 0% of tweets were neutral, 28.7% were positive, and 71.3% were negative. Figure 6 depicts a time graph from 3 August to 6 August 2021, plotted against the number of tweets and retweets. The recorded time started at 12:00 am and ended at 11:59 pm. The graph shows that the highest time for tweets during daylight is between 11:00 am and 1:59 pm. It also recorded a higher number of tweets and retweets, with 46.17% of the highest tweets. Meanwhile, during nighttime, tweeps were active between 10:00 pm and 2:59 am.
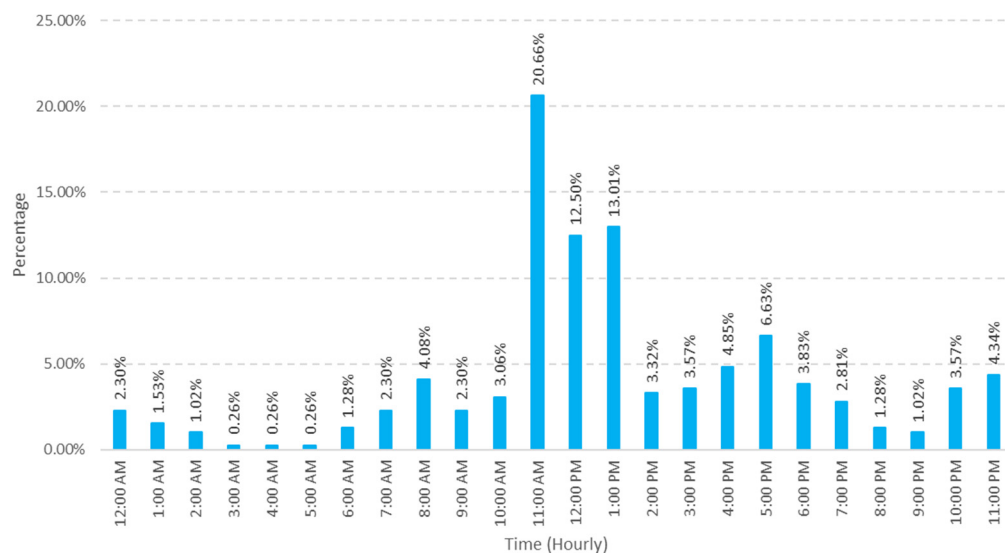

Figure 6. Numbers of tweets hourly from 3 August to 6 August 2021.

We explain the timing of the tweets using the sentiment score graph (Figure 7) from 3 August to 6 August 2021. The graph shows the range of positive (compound ≥ 0.05), neutral (compound > -0.05, and compound < 0.05), and negative (compound ≤ -0.05). To determine the percentage of positive, negative, and neutral sentiment, we counted the number of tweets in each category and divided it by the total number of tweets, then multiplied by 100. Figure 7 shows that 71.3% of tweets are classified as postings with negative sentiments.
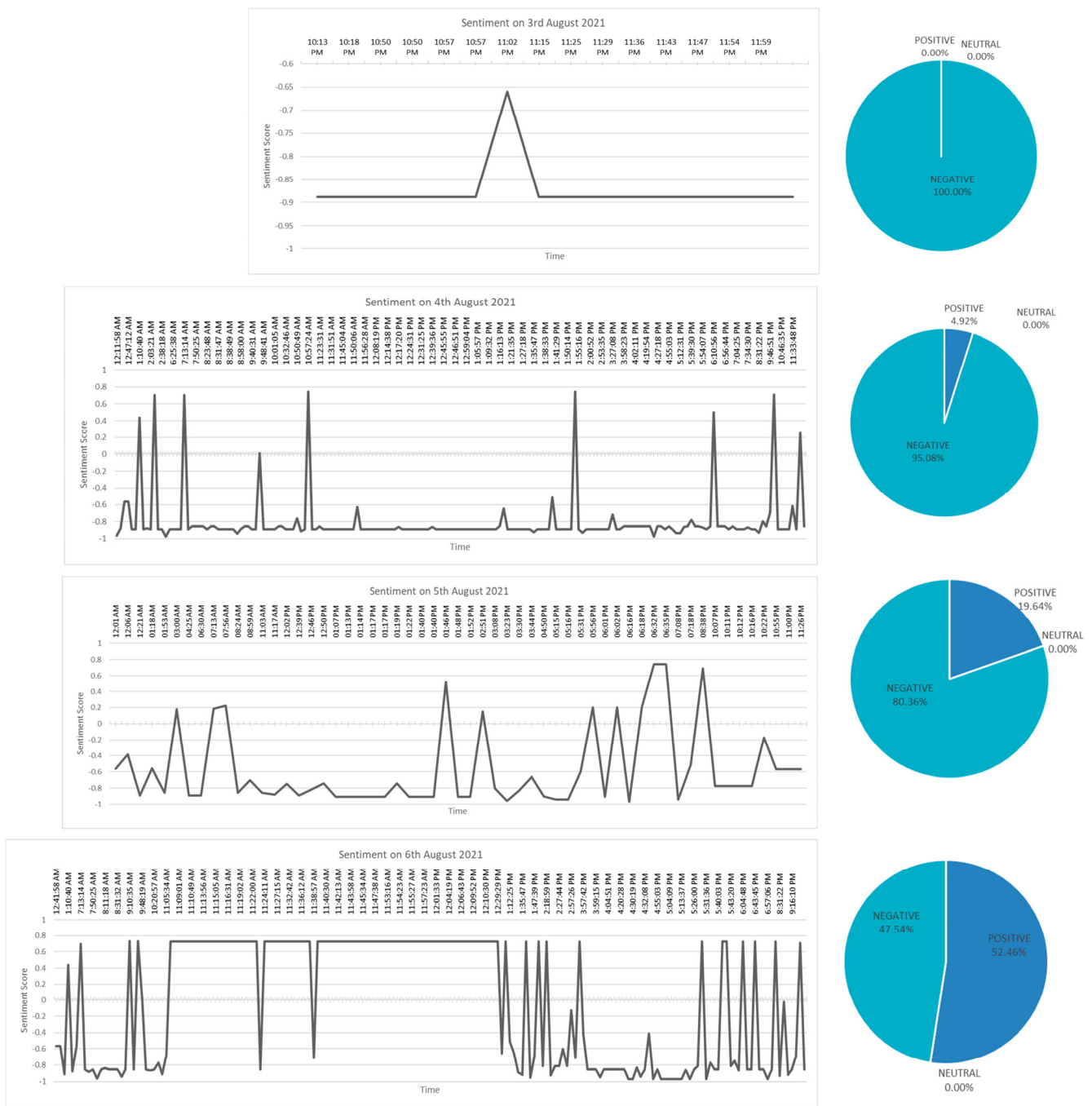
Figure 7. A sentiment value of tweets from 3 August to 6 August 2021.

However, there were also times when tweeps posted positive sentiments (28.7% of tweets). A positive sentiment (0.7269) was shown on 6 August from 11:00 am to 8:00 pm, which was a tweet that originated from Kangar (a city in the state of Perlis, north Malaysia) and retweeted many times. Although the tweet appeared to have a pleasant message, it was actually sarcastic and mocked the authorities. It began with a thank you to the Ministry of Health (known as KKM), which finally confirmed that the manufacturing sector was to blame for the rise of COVID-19 cases. The truth was that the government was warned about it for a very long time, and the public was aware of it. Unfortunately, the government was adamant about restricting people's freedom of movement while allowing the manufacturing sector to operate (refer to Table 2).

The subsequent tweet also exudes a saccharine atmosphere by detailing the government's "achievements" in dealing with COVID-19 in Malaysia. This tweet was sent at 6:32 pm on 5 August with a sentiment score of 0.743.

"Success story:

- COVID 16,000+ per day
- Death due to COVID 200+ per day
- 1000+ suicide
- The nation's poverty rate increased by double.
- More than 500,000 M40 families currently hold the status of B40.
- 1/4 graduates unemployed today!"

Malaysians were also enraged by the Prime Minister's efforts to appoint new ministers to address Malaysia's economic problems. Due to the large number of ministers in the current government cabinet (32 ministers and 38 deputy ministers), this nomination was exceedingly unpopular with the populace, who felt that existing ministers should be able to address most national problems without adding a new minister. The appointment of additional ministers was a controversial issue because it was seen as a way for the government to increase its size and cost without necessarily improving its effectiveness. It is important for the government to consider the views and concerns of the public when making decisions and to try to communicate their reasoning and rationale to the public transparently and clearly.

The following tweet (5 August 2020, 5:56 pm) has a sentiment score of 0.2023:

"The approach proposed by Prime Minister Mahiadin is appropriate for escaping political concerns, the economy, the COVID-19 pandemic, suicides, hunger, and job and income losses. He was boosting ministry members' responsibilities and revenue."

This means positive comments were not necessarily related to positive scores through sentiment analysis because the value might differ when interpreted in different contexts and by understanding the community. As a result, we must consider this fact when applying this analysis, particularly in cases of other cultural relativism, such as communities that engage in sarcasm. Meanwhile, other tweets conveying positive values were associated with the tweeters' positive attitudes, projecting their psychological status onto them. It is where people express their positive feelings by using constructive or intensifying positive words in the text. The following tweet has a sentiment score of 0.7096:

"To those who have given up, please be rest assured that Allah's help is so near. Don't commit suicide. Allah will help us."

It appears to encourage people who may be discouraged or feel hopeless to persevere and have faith that things will improve. It emphasized that help was near and encouraged people not to give up or take drastic measures such as committing suicide. It is important to remember that it is okay to feel down or struggle at times, and resources and support are available for people struggling with their mental health or feeling overwhelmed.

A total of 100% of the tweets on 3 August 2020 (Figure 7) contributed to negative sentiment. Tweeps were most active between 10:00 pm and 11:59 pm. It also shows that the highest frequency of tweet occurrences creates a cycle of social activity among the tweeps. An interesting point to note here is that the frequency of negative sentiment scores at night was higher than negative sentiment scores during the day. The negative sentiment scores show that Malaysians were in distress and actively complained and expressed their irritation and disgust. It was loaded with negative feelings of anger, frustration, worry, dissatisfaction, and hopelessness that may lead to nighttime anxiety. As a result, this phenomenon causes most tweeps to stay up at night and continue to express their feelings through tweets, potentially leading to insomnia that changes their daily circadian rhythm. This is in line with the findings of Pappa [9], who stated that negative feelings that keep one awake all night might lead to nighttime anxiety and, ultimately, insomnia.

From 10:00 pm to 11:59 pm on 3 August 2020, the following tweet expressed an unfavorable sentiment with a score of −0.8883:

"I understand if there are still T20 classes supporting the government. Maybe they have their own benefits that need to be taken care of. But why are the B40 classes still supporting them? Don't you guys see the destruction happening in this country?"

The tweet apparently expressed a negative sentiment towards the government and questioned why the B40 class (a term used in Malaysia to refer to the bottom 40% of households in terms of income) was still supporting them. It suggested that the government was destroying the country. The top 20% of earners are designated as T20, which typically consists of high-earners, including executives, professionals, and business owners. On the other hand, the B40 group consists of those who make less money, e.g., those who work in positions that pay less or who are unemployed or underemployed. The pandemic likely disproportionately impacted lower income households, including the B40 class in Malaysia. The economic downturn caused by the pandemic and measures taken to control the spread of the virus, such as lockdowns and business closures, may have had a greater impact on lower income households, which may have limited financial resources to fall back on. This could widen the gap between lower income households and higher income households.

It is typical for people to hold varying political viewpoints and beliefs. Political parties' and administrations' goals and priorities frequently differ, so what may be advantageous for one individual or group may not be for another. It is also likely that some people will continue to support the administration because they believe that measures have been taken to help them or meet their requirements. Political matters are complicated, and there are frequently many variables at work. It is not always simple to fully comprehend why someone might have a particular belief or make a particular choice.

Figure 8 provides a visual representation of the location of the sentiment analysis. As a result of this study, we found a way to use ArcGIS to visualize the data mined from Twitter. This particular component of our study is what distinguishes it from others. ArcGIS is a popular GIS software that is used to create and display maps, as well as to analyze and manage spatial data. Visualizing data using ArcGIS helps in understanding trends and patterns in the data. In addition, it can help to make the results more accessible and easier to understand public opinion and attitudes towards different issues.
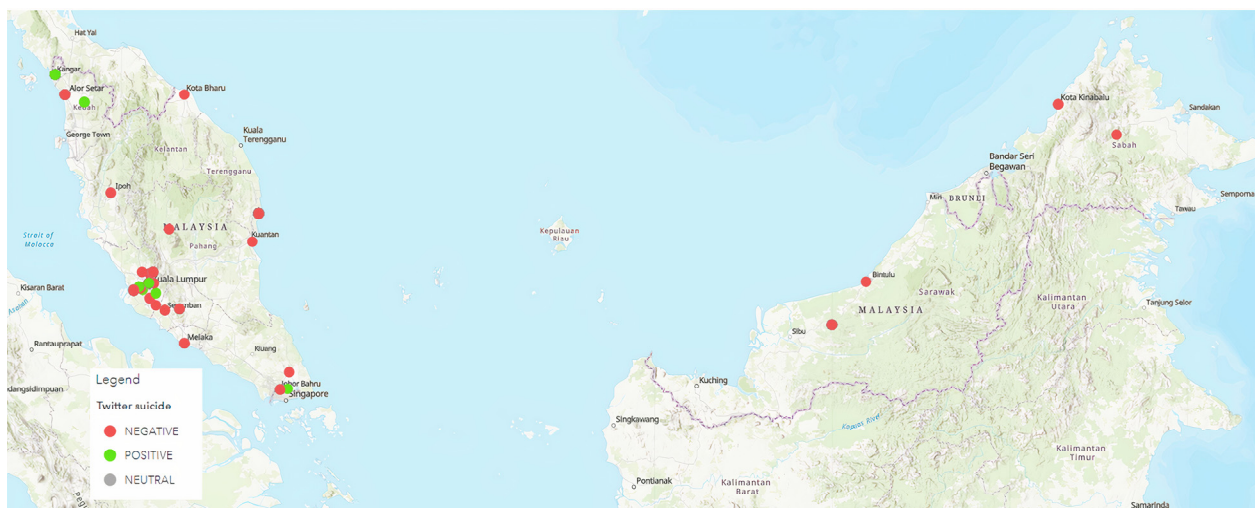

Figure 8. Sentiment analysis displayed spatially.

The distribution of green (positive) and red (negative) across Malaysia shows a pattern of both positive and negative emotions (Figure 8), with a negative sentiment being the more common of the two. Based on the sentiment-by-place distribution, we discovered there were 392 fewer tweets than anticipated, most likely because many tweets were forwarded numerous times. However, it is impossible to determine the location of the individuals who forwarded the tweet. Only the original user's location can be determined, which may lead to fewer tweets associated with certain locales.

Despite the fact that not all Twitter data can be used to indicate the location of sentiment, this study successfully proved a highly useful output since it can reveal geographic disparities in public opinion or sentiment, especially for tweets related to suicide. It can help pinpoint areas with an unusually strong or weak sentiment. It can help locate potential regions or hotspot problems, such as places with a high concentration of unfavorable sentiment.

4. Conclusions

This study aimed to examine social media data from Twitter to understand the context of how Malaysians used the word "suicide" during the COVID-19 outbreak. Our findings include a few significant conclusions. First, according to the spatial pattern distribution, Kuala Lumpur, Klang, Subang Jaya, Kangar, Alor Setar, Chukai, Kuantan, Johor Bharu, and Kota Kinabalu were cities that sent the most tweets. This data gives the sentiment analysis of the term "suicide" on Twitter in a geographic context and can be used to spot regional variations in the word's usage.

Second, the content analysis conducted on the tweets related to the keyword "suicide" revealed three main themes: (i) criticism of the government of that day, (ii) awareness related to suicide, and (iii) suicide feelings or experiences. Malaysian tweeps used the keyword to express both negative and positive sentiments. They posted negative tweets in response to the government's handling of the COVID-19 pandemic and its impacts on the economy and citizens. The public expressed concern over the increasing number of COVID-19 cases and its impact on the economy, leading to a higher rate of unemployment and poverty and its correlation to the rise in suicide cases. Political instability added to the frustration and disappointment with the government, highlighting the emotional frustration and dissent of the public. Paradoxically, the keyword "suicide" also conveyed positive messages of support for mental health and relationship-building. Some tweets encouraged positivity and understanding and offered assistance. Malaysians also spread hope, encouragement, and support, reminding people that help was always available. This shows that the use of the word "suicide" varied based on the local context and social culture.

In the Malaysian context, it is important to note that sentiment analysis may not always reflect the true sentiment due to sarcasm, cultural differences, and differences in interpretation. This study, therefore, highlights the need for caution when interpreting the results of sentiment analysis and the importance of considering cultural and regional differences. In future, this information can be used to identify areas that may require additional resources or interventions to address suicide risk and to obtain a more complete and nuanced understanding of the phenomenon of suicide and its social and cultural context.

This study also sheds light on the potential of using social media big data to understand the virtual behavior of the public in an unprecedented crisis. The pandemic has highlighted the importance of real-time information and the need for quick and effective decision making. Social media platforms generate large amounts of real-time data that we can analyze for valuable insights. Data mining is a cost-effective method of gathering information compared to traditional data collection methods, providing researchers with a wealth of information without the need for complex and time-consuming data collection processes.

Social media provides an unfiltered platform for people to express their opinions and sentiments, providing valuable insights into public opinion and attitudes toward a specific topic. The results of this study show the potential of social media data to formulate public health policies and to identify areas that require additional resources or interventions. In the context of the COVID-19 pandemic, we used social media data to monitor the impact of the pandemic on mental health and to identify potential risk factors for suicide. By using social media data, policymakers and public health professionals can respond quickly and effectively to the population's needs and address public health concerns in real time.

This study has crucial implications for Malaysian public health policies and actions during the COVID-19 pandemic. Policymakers and public health professionals can acquire significant insights into the public's opinions and sentiments around suicide, mental health, and government responses to the pandemic by examining social media data. The study underlines the need to take cultural and geographical differences into account when assessing sentiment data and understand the local environment in molding public opinion. Our findings further emphasize the value of social media big data in analyzing virtual behavior and making fast judgments. Social media networks create massive volumes of real-time data that may be evaluated for important insights, offering a cost-effective alternative to sophisticated and time-consuming data collection techniques. Policymakers and public health experts may respond quickly and effectively to population needs and address public health concerns in real time by leveraging social media data.

While our study focuses on Malaysians' views and sentiments regarding suicide during the COVID-19 pandemic, the findings may also have significance for urban planning. The study shows regional trends in the distribution of tweets about suicide, highlighting the geographic context in which the phrase was used. This data could help city planners discover places with a higher prevalence of suicide-related tweets, suggesting a higher risk of mental health difficulties in these areas. Furthermore, we found that how the government managed the pandemic and its impact on the economy and citizens contributed to suicidal feelings. Consequently, this points to a need for governmental initiatives to improve urban economic and social conditions. The role of urban planning in providing livable, healthy, and sustainable environments that support citizens' mental health and well-being is critical.

In conclusion, this study provides valuable insights into Malaysians' attitudes and sentiments toward the keyword "suicide" during the pandemic. The themes we identified in the tweets provide a unique perspective on the challenges Malaysians faced during this difficult time and the role that social media can play in spreading awareness and supporting mental health and suicide prevention. The study highlights the importance of considering the local context and cultural differences when analyzing social media data and the potential of social media big data for understanding virtual behavior and making timely decisions.

References

1. Honey-Rosés, J.; Anguelovski, I.; Chireh, V.K.; Daher, C.; van den Bosch, C.K.; Litt, J.S.; Mawani, V.; McCall, M.K.; Orellana, A.; Oscilowicz, E.; et al. The impact of COVID-19 on public space: An early review of the emerging questions—design, perceptions and inequities. *Cities Health* **2021**, *5*, S263–S279. [CrossRef]
2. Simon, D.; Arano, A.; Cammisa, M.; Perry, B.; Pettersson, S.; Riise, J.; Valencia, S.; Oloko, M.; Sharma, T.; Vora, Y.; et al. Cities coping with COVID-19: Comparative perspectives. *City* **2021**, *25*, 129–170. [CrossRef]
3. Mehta, V. The new proxemics: COVID-19, social distancing, and sociable space. *J. Urban Des.* **2020**, *25*, 669–674. [CrossRef]
4. Goldstein, J.R.; Lee, R.D. Demographic perspectives on the mortality of COVID-19 and other epidemics. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 22035–22041. [CrossRef]
5. Ravindran, T. When a pandemic intensifies racial terror: The politics of COVID-19 control in Bolivia. *City* **2020**, *24*, 778–792. [CrossRef]
6. Humphry, D. Campaigning in the time of coronavirus. *City* **2020**, *24*, 423–430. [CrossRef]
7. Carandang, B.; Tobias, L.-A. How Fintech Can Help South-East Asia's MSMEs Recover from COVID-19. 2020. Available online: https://www.weforum.org/agenda/2020/07/how-fintech-can-help-south-east-asias-msmes-recover-from-covid-19/ (accessed on 5 December 2022).
8. Balakrishnan, V.; Ng, K.S.; Kaur, W.; Lee, Z.L. COVID-19 mental health prevalence and its risk factors in South East Asia. *Curr. Psychol.* **2022**, 1–16. [CrossRef]
9. Pappa, S.; Chen, J.; Barnett, J.; Chang, A.; Dong, R.K.; Xu, W.; Yin, A.; Chen, B.Z.; Delios, A.Y.; Chen, R.Z.; et al. A systematic review and meta-analysis of the mental health symptoms during the COVID-19 pandemic in Southeast Asia. *Psychiatry Clin. Neurosci.* **2021**, *76*, 41–50. [CrossRef]
10. World Health Organization. Mental Health and Psychosocial Considerations during the COVID-19 Outbreak, 18 March 2020. Available online: https://apps.who.int/iris/handle/10665/331490 (accessed on 20 November 2022).
11. Osborn, D.P.; King, M.; Nazareth, I.; Levy, G.; Petersen, I. Suicide and COVID-19: A rapid systematic review. *J. Affect. Disord.* **2020**, *275*, 535–542.
12. World Health Organization. Suicide. 2019. Available online: https://www.who.int/mental_health/suicide-prevention/suicide/en/ (accessed on 19 November 2022).
13. Turecki, G.; Brent, D.A.; Gunnell, D.; O'Connor, R.C.; Oquendo, M.A.; Pirkis, J.; Stanley, B.H. Suicide and suicide risk. *Nat. Rev. Dis. Prim.* **2019**, *5*, 1–22. [CrossRef]
14. American Psychological Association. *Suicide and Suicide Prevention*; American Psychological Association: Washington, DC, USA, 2022; Available online: https://www.apa.org/topics/suicide (accessed on 15 January 2023).
15. Mars, B.; Heron, J.; Klonsky, E.D.; Moran, P.; O'connor, R.C.; Tilling, K.; Wilkinson, P.; Gunnell, D. Predictors of future suicide attempt among adolescents with suicidal thoughts or nonsuicidal self-harm: A population-based birth cohort study. *Lancet Psychiatry* **2019**, *6*, 327–337. [CrossRef] [PubMed]
16. Klonsky, E.D.; May, A.M.; Saffer, B.Y. Suicide, suicide attempts, and suicidal ideation. *Annu. Rev. Clin. Psychol.* **2016**, *12*, 307–330. [CrossRef] [PubMed]
17. Sarker, A.; O'Connor, K.; Ginn, R.; Scotch, M.; Smith, K.; Malone, D.; Gonzalez, G. Social media mining for toxicovigilance: Automatic monitoring of prescription medication abuse from twitter. *Drug Saf.* **2016**, *39*, 231–240. [CrossRef] [PubMed]
18. Fergusson, L.J.; Woodward, D.M.; Horwood, L.J. Risk factors and life processes associated with the onset of suicidal behaviour during adolescence and early adulthood. *Psychol. Med.* **2000**, *30*, 23–39. [CrossRef] [PubMed]
19. Bentley, K.H.; Franklin, J.C.; Ribeiro, J.D.; Kleiman, E.M.; Fox, K.R.; Nock, M.K. Anxiety and its disorders as risk factors for suicidal thoughts and behaviors: A meta-analytic review. *Clin. Psychol. Rev.* **2015**, *43*, 30–46. [CrossRef] [PubMed]
20. Brent, D.A.; Oquendo, M.; Birmaher, B.; Greenhill, L.; Kolko, D.; Stanley, B.; Zelazny, J.; Brodsky, B.; Melhem, N.; Ellis, S.P.; et al. Familial transmission of mood disorders: Convergence and divergence with transmission of suicidal behavior. *J. Am. Acad. Child Adolesc. Psychiatry* **2004**, *43*, 1259–1266. [CrossRef] [PubMed]
21. Kim, C.D.; Seguin, M.; Therrien, N.; Riopel, G.; Chawky, N.; Lesage, A.D.; Turecki, G. Familial aggregation of suicidal behavior: A family study of male suicide completers from the general population. *J. Psychiatry* **2005**, *162*, 1017–1019. [CrossRef]
22. Brent, D.A.; Melhem, N.M.; Oquendo, M.; Burke, A.; Birmaher, B.; Stanley, B.; Biernesser, C.; Keilp, J.; Kolko, D.; Ellis, S.; et al. Familial pathways to early-onset suicide attempt: A 5.6-year prospective study. *JAMA Psychiatry* **2015**, *72*, 160–168. [CrossRef]
23. Burrell, L.V.; Mehlum, L.; Qin, P. Sudden parental death from external causes and risk of suicide in the bereaved offspring: A national study. *J. Psychiatr. Res.* **2018**, *96*, 49–56. [CrossRef]
24. Lutz, P.E.; Mechawar, N.; Turecki, G. Neuropathology of suicide: Recent findings and future directions. *Mol. Psychiatry* **2017**, *22*, 1395–1412. [CrossRef]
25. Jollant, F.; Bellivier, F.; Leboyer, M.; Courtet, P. Suicide in affective disorders: A critical review of the literature. *J. Affect. Disord.* **2017**, *207*, 63–74.
26. Nock, M.K.; Green, J.G.; Hwang, I.; McLaughlin, A.M.; Sampson, K.A.; Zaslavsky, N.A.; Kessler, R.C. Prevalence, correlates, and treatment of lifetime suicidal behavior among adolescents: Results from the National Comorbidity Survey Replication Adolescent Supplement. *J. Clin. Psychiatry* **2015**, *76*, 1023–1032. [CrossRef] [PubMed]
27. El-Mohandes, E.; Mohandes, A.A. Post-traumatic stress disorder and suicide: A systematic review. *J. Clin. Psychol.* **2018**, *74*, 945–966.

28. CodeBlue. Malaysia Records Three Daily Suicides This Year up to May 2021. Available online: https://codeblue.galencentre.org/2021/07/01/malaysia-records-three-daily-suicides-this-year-up-to-may/ (accessed on 1 May 2021).

29. Ministry of Health Malaysia. COVIDNOW in Malaysia. 2022. Available online: https://covidnow.moh.gov.my/ (accessed on 13 May 2021).

30. Gandomi, A.; Haider, M. Beyond the hype: Big data concepts, methods, and analytics. *Int. J. Inf. Manag.* **2015**, *35*, 137–144. [CrossRef]

31. Benfeldt, O.; Persson, J.S.; Madsen, S. Data Governance as a Collective Action Problem. *Inf. Syst. Front.* **2020**, *22*, 299–313. [CrossRef]

32. Abraham, R.; Schneider, J.; vom Brocke, J. Data governance: A conceptual framework, structured review, and research agenda. *Int. J. Inf. Manag.* **2019**, *49*, 424–438. [CrossRef]

33. Al-Ruithe, M.; Benkhelifa, E.; Hameed, K. A systematic literature review of data governance and cloud data governance. *Pers. Ubiquitous Comput.* **2019**, *23*, 839–859. [CrossRef]

34. Digital Marketing Institute. Social Media: What Countries Use It Most & What Are They Using. 2021. Available online: https://digitalmarketinginstitute.com/blog/social-media-what-countries-use-it-most-and-what-are-they-using (accessed on 1 April 2022).

35. Müller, J. Active Social Media Users as Percentage of the Total Population in Malaysia from 2016 to 2021. Statista 2021. Available online: https://www.statista.com/statistics/883712/malaysia-social-media-penetration/ (accessed on 1 April 2022).

36. Wong, A.; Ho, S.; Olusanya, O.; Antonini, M.V.; Lyness, D. The use of social media and online communications in times of pandemic COVID-19. *J. Intensive Care Soc.* **2021**, *22*, 255–260. [CrossRef]

37. Blanchard, A. Virtual Behavior Settings: An Application of Behavior Setting Theories to Virtual Communities. *J. Comput. -Mediat. Commun.* **2004**, *9*, JCMC924. [CrossRef]

38. Musa, K.; Arifin, W.; Mohd, M.; Jamiluddin, M.; Ahmad, N.; Chen, X.; Hanis, T.; Bulgiba, A. Measuring time-varying effective reproduction numbers for COVID-19 and their relationship with movement control order in Malaysia. *Int. J. Environ. Res. Public Health* **2021**, *18*, 3273. [CrossRef]

39. Saud, M.; Mashud, M.; Ida, R. Usage of social media during the pandemic: Seeking support and awareness about COVID-19 through social media platforms. *J. Public Aff.* **2020**, *20*, e2417. [CrossRef]

40. Nadeak, B. The effectiveness of distance learning using social media during the pandemic period of COVID-19: A case in universitas kristen indonesia. *Int. J. Adv. Sci. Technol.* **2020**, *29*, 1764–1772.

41. QSR International Pty Ltd. NVivo (Version 12). Available online: https://lumivero.com/products/nvivo/ (accessed on 15 July 2022).

42. Jackson, K.; Bazeley, P. *Qualitative Data Analysis with NVivo*, 3rd ed.; SAGE Publications: Los Angeles, CA, USA, 2019.

43. García-Ramírez, G.M.; Bogen, K.W.; Rodríguez-Guzmán, V.M.; Nugent, N.; Orchowski, L.M. #4645Boricuas: Twitter reactions to the estimates of deaths by Hurricane María in Puerto Rico. *J. Community Psychol.* **2021**, *49*, 768–790. [CrossRef] [PubMed]

44. Bogen, K.W.; Millman, C.; Huntington, F.; Orchowski, L.M. A Qualitative Analysis of Disclosing Sexual Victimization by #NotOkay during the 2016 Presidential Election. *Violence Gend.* **2018**, *5*, 174–181. [CrossRef]

45. Twitter. Twitter Developers: Geo Objects. 2019. Available online: https://developer.twitter.com/en/docs/tweets/data-dictionary/overview/geo-objects (accessed on 15 October 2022).

46. Lachmar, E.M.; Wittenborn, A.K.; Bogen, K.W.; McCauley, H.L. #MyDepressionLooksLike: Examining public discourseabout depression on Twitter. *J. Med. Internet Res. Ment. Health* **2017**, *4*, e43. [CrossRef]

47. Twitter. How to Change Your Country Setting? *2022.* Available online: https://help.twitter.com/en/managing-your-account/how-to-change-country-settings (accessed on 1 April 2022).

48. Alotaibi, A.; Rahman, A.-U.; Alhaza, R.; Alkhalifa, W.; Alhajjaj, N.; Alharthi, A.; Abushoumi, D.; Alqahtani, M.; Alkhulaifi, D. Spam and Sentiment Detection in Arabic Tweets Using MarBert Model. *Int. Inf. Eng. Technol. Assoc.* **2022**, *9*, 1574–1582. [CrossRef]

49. Alqarni, A.; Rahman, A. Arabic Tweets-Based Sentiment Analysis to Investigate the Impact of COVID-19 in KSA: A Deep Learning Approach. *Big Data Cogn. Comput.* **2023**, *7*, 16. [CrossRef]

50. Steven, B.; Loper, E.; Klein, E. *Natural Language Processing with Python*; O'Reilly Media Inc.: Sebastopol, CA, USA, 2009.

51. Hutto, C.J.; Gilbert, E.E. VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. In Proceedings of the Eighth International Conference on Weblogs and Social Media (ICWSM-14), Ann Arbor, MI, USA, 1–4 June 2014.

52. Çilgin, C.; Baş, M.; Bilgehan, H.; Ünal, C. Twitter Sentiment Analysis During COVID-19 Outbreak with VADER. *Acad. J. Inf. Technol.* **2022**, *13*, 72–89. [CrossRef]

53. Elbagir, S.; Yang, J. Twitter Sentiment Analysis Using Natural Language Toolkit and VADER Sentiment. In Proceedings of the International MultiConference of Engineers and Computer Scientists, Hong Kong, China, 13–15 March 2019; pp. 122–160.

54. Sokolova, M.; Lapalme, G. A systematic analysis of performance measures for classification tasks. *Inf. Process. Manag.* **2009**, *45*, 427–437. [CrossRef]

55. Kiritchenko, S.; Mohammad, S.M. Capturing Reliable Fine-Grained Sentiment Associations by Crowdsourcing and Best–Worst Scaling. In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, San Diego, CA, USA, 12–17 June 2016; pp. 811–817. [CrossRef]

56. Tausczik, Y.R.; Pennebaker, J.W. The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *J. Lan Soc. Psychol.* **2010**, *29*, 24–54. [CrossRef]

57.　Read, J.; Carroll, J. Weakly supervised techniques for domain-independent sentiment classification. In Proceedings of the 1st International CIKM Workshop on Topic-Sentiment Analysis for Mass Opinion, Hong Kong, China, 6 November 2009; pp. 45–52. [CrossRef]

58.　Hsieh, H.-F.; Shannon, S.E. Three Approaches to Qualitative Content Analysis. *Qual. Health Res.* **2005**, *15*, 1277–1288. [CrossRef] [PubMed]

59.　Bengtsson, M. How to plan and perform a qualitative study using content analysis. *Nurs. Open* **2016**, *2*, 8–14. [CrossRef]

60.　Senz, K. Outrage Spreads Faster on Twitter: Evidence from 44 News Outlets. Harvard Business School Working Knowledge. 2021. Available online: https://hbswk.hbs.edu/item/hate-spreads-faster-on-twitter-evidence-from-44-news-outlets (accessed on 11 February 2023).

61.　Department of Statistics Malaysia. Labour Force Survey Report, Malaysia, 2019. 2020. Available online: https://www.dosm.gov.my/v1/index.php?r=column/pdfPrev&id=TlVMbEtBVXBGTi80VjdqZ1JUdVRHdz09 (accessed on 18 December 2022).

62.　Department of Statistics Malaysia. Labour Force Survey Report, Malaysia 2020. 2021. Available online: https://www.dosm.gov.my/v1/index.php?r=column/cthemeByCat&cat=126&bul_id=dTF2dkJpcUFYUWRrczhqUHVpcDRGQT09&menu_id=Tm8zcnRjdVRNWWlpWjRlbmtlaDk1UT09 (accessed on 18 December 2022).

63.　Department of Statistics Malaysia. Principal Statistics of Labour Force, Malaysia, Third Quarter (Q3) 2022. 2022. Available online: https://www.dosm.gov.my/v1/index.php?r=column/cthemeByCat&cat=149&bul_id=bjBadGlBL2JHaFpZdkZKeHc3eUIyQT09&menu_id=Tm8zcnRjdVRNWWlpWjRlbmtlaDk1UT09#:~:text=Principal%20Statistics%20of%20Labour%20Force,%2C%20Third%20Quarter%20(Q3)%202022&text=In%20the%20third%20quarter%20of,persons%20registering%2016.44%20million%20persons (accessed on 18 December 2022).

64.　Black, P.E. Euclidean Distance. In Dictionary of Algorithms and Data Structures. 2004. Available online: https://www.nist.gov/dads/HTML/euclidndstnc.html (accessed on 17 January 2023).

65.　Department of Statistic Malaysia (DOSM). Population by State. 2021. Available online: https://www.dosm.gov.my/v1/index.php?r=column/cthemeByCat&cat=155&bul_id=ZjJOSnpJR21sQWVUcUp6ODRudm5JZz09&menu_id=L0pheU43NWJwRWVSZklWdzQ4TlhUUT09#:~:text=Kuala%20Lumpur%20had%20the%20highest,1%2C691%20people)%20per%20square%20kilometre (accessed on 12 October 2022).