



Contents lists available at ScienceDirect

# International Journal of Applied Earth Observations and Geoinformation

journal homepage: [www.elsevier.com/locate/jag](http://www.elsevier.com/locate/jag)

## A global portrait of expressed mental health signals towards COVID-19 in social media space

Siqin Wang<sup>a,j,1,\*</sup>, Xiao Huang<sup>b,j,\*,1</sup>, Tao Hu<sup>c,j,1</sup>, Bing She<sup>d,1</sup>, Mengxi Zhang<sup>e,j</sup>, Ruomei Wang<sup>a</sup>, Oliver Gruebner<sup>f,g</sup>, Muhammad Imran<sup>h</sup>, Jonathan Corcoran<sup>a</sup>, Yan Liu<sup>a</sup>, Shuming Bao<sup>i</sup>

<sup>a</sup> School of Earth and Environmental Sciences, University of Queensland, Brisbane, Queensland, Australia

<sup>b</sup> Department of Geosciences, University of Arkansas, AR, USA

<sup>c</sup> Department of Geography, Oklahoma State University, OK, USA

<sup>d</sup> Institute for Social Research, University of Michigan, MI, USA

<sup>e</sup> Department of Nutrition and Health Science, Ball State University, IN, USA

<sup>f</sup> Department of Geography, University of Zurich, Switzerland

<sup>g</sup> Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Switzerland

<sup>h</sup> Qatar Computing Research Institute, Hamad Bin Khalifa University, Qatar

<sup>i</sup> China Data Institute and Future Data Lab, MI, USA

<sup>j</sup> Spatial Data Lab, Centre for Geographic Analysis, Harvard University, MA, USA

### ARTICLE INFO

#### Keywords:

Negative sentiments  
Multilingual tweets  
Sentiment analysis  
Social media  
Policy implementation  
COVID-19  
Pandemic

### ABSTRACT

Globally, the COVID-19 pandemic has induced a mental health crisis. Social media data offer a unique opportunity to track the mental health signals of a given population and quantify their negativity towards COVID-19. To date, however, we know little about how negative sentiments differ across countries and how these relate to the shifting policy landscape experienced through the pandemic. Using 2.1 billion individual-level geotagged tweets posted between 1 February 2020 and 31 March 2021, we track, monitor and map the shifts in negativity across 217 countries and unpack its relationship with COVID-19 policies. Findings reveal that there are important geographic, demographic, and socioeconomic disparities of negativity across continents, different levels of a nation's income, population density, and the level of COVID-19 infection. Countries with more stringent policies were associated with lower levels of negativity, a relationship that weakened in later phases of the pandemic. This study provides the first global and multilingual evaluation of the public's real-time mental health signals to COVID-19 at a large spatial and temporal scale. We offer an empirical framework to monitor mental health signals globally, helping international authorizations, including the United Nations and World Health Organization, to design smart country-specific mental health initiatives in response to the ongoing pandemic and future public emergencies.

### 1. Introduction

The COVID-19 pandemic has caused profound impacts on our everyday lives, threatening our mental health (Cullen et al., 2020; John F. Helliwell, 2021). Our negative sentiments (e.g., fear, insecurity, and anxiety) expressed towards COVID-19 have been observed across a number of countries, including the United States (U.S.) (Czeisler et al., 2020; Hu et al., 2021; Jacobson et al., 2020), the United Kingdom (U.K.) (Carr et al., 2021; O'Connor et al., 2021), Australia (Ewing and Vu, 2021; Fisher et al., 2020; Newby et al., 2020; Tan et al., 2020; Tran et al.,

2020; Van Rheenen et al., 2020; Wang et al., 2022), China (Liu et al., 2020; Ren et al., 2020; Talevi et al., 2020; Wang et al., 2021), and European nations (Hummel et al., 2021). Negative sentiments are principally associated with the fear of COVID-19 infection, the financial burdens related to business closures, reduced economic activities, and the government mobility restriction policies designed to control viral spread (Jacobson et al., 2020; Wang et al., 2021). The World Happiness Report published by the United Nations concluded that the threat to the mental health and happiness of human beings due to COVID-19 necessitates an urgent response (John F. Helliwell, 2021). Timely measuring,

\* Corresponding authors.

E-mail addresses: [s.wang6@uq.edu.au](mailto:s.wang6@uq.edu.au) (S. Wang), [xh010@uark.edu](mailto:xh010@uark.edu) (X. Huang).

<sup>1</sup> These authors contribute equally to this research.

<https://doi.org/10.1016/j.jag.2022.103160>

Received 7 September 2022; Received in revised form 7 November 2022; Accepted 15 December 2022

Available online 17 December 2022

1569-8432/© 2022 Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

mapping, and monitoring of the mental health signals of human society represent an urgent task to cope with the emerging mental health crisis—that is, to quantify the public’s negative sentiments towards COVID-19 as a litmus test to the mental health of a nation and to unveil the scope, scale and characteristics of the mental health crisis.

Scholarship on mental health and happiness has traditionally employed survey-based assessments (Balcombe and De Leo, 2020), including those examining the COVID-19 context (Ewing and Vu, 2021; Fisher et al., 2020; Newby et al., 2020; Tan et al., 2020; Tran et al., 2020; Van Rheenen et al., 2020). However, these studies are limited to capture long-term trends alongside real-time responses to COVID-19 and are unable to measure mental health signals at scale (Balcombe and De Leo, 2020). Alternatively, social media data (e.g., Twitter) offers a unique opportunity to address some of the limitations associated with survey-based sources and measure the sentiment of populations in the face of certain circumstances (Hu et al., 2021; Hussain et al., 2021; Wang et al., 2020; Wang et al., 2022). The text-based contents (e.g., expressions, words, and languages,) expressed by social media users can offer a unique indicator of an individual’s emotional response to a particular event or phenomenon (Agarwal et al., 2011; Kouloumpis et al., 2011). Further, the individual signals of sentiment and emotion provide critical insights into the mental health status of aggregated populations (Coppersmith et al., 2014). A large body of studies using social media data usually rely on advanced techniques (e.g., artificial intelligent models, machine and deep learning algorithms) to quantify sentiment in reflection of mental health (Ewing and Vu, 2021; Kwok et al., 2021; Zhou et al., 2021). However, this type of study predominantly focuses on single language contents, one or several countries or regions, and the data within a short time period, especially at the early stage of the pandemic. Building on a growing body of studies using social media data to examine mental health signals, we scale our empirical framework to examine the global and multilingual context to unpack the way in which the COVID-19 pandemic has affected the mental health of human societies.

Drawing on 2.1 billion individual-level geotagged tweets posted from 1 February 2020 to 31 March 2021 covering 217 countries and 67 languages, we address three research questions related to negative sentiments (hereinafter termed as negativity) towards COVID-19: 1) How does negativity shift over time and vary across geographic, demographic and socioeconomic contexts?; 2) Are there any temporal clusters of negativity across countries and along the pandemic timeline?; and, 3) To what extent is the negativity associated with shifts in policy

implementations? This study contributes the global multilingual evaluation of mental health signals in relation to COVID-19, and the analytical framework is designed in a manner such that it can be deployed to capture real-time mental health signals responding to the ongoing pandemic and future public health emergencies.

## 2. Materials and methods

### 2.1. Tweet sentiment data

The research design and analytical workflow of our study was shown in Fig. 1. We used large-scale Twitter data containing around two billion multilingual tweets related to the COVID-19 pandemic from 87 million unique users worldwide (Imran et al., 2022). The dataset was acquired through the Twitter Streaming API (Twitter, 2020), using + 800 multilingual keywords (Supplementary Table S8) relevant to a large set of topics, including COVID-19 policies and cases (e.g., mask usage, social distancing and infection symptoms). The dataset’s large spatial coverage (i.e., tweets from 217 countries) and temporal span (i.e., data collection from 1 February 2020 to 31 March 2021) offer unique opportunities to study the public’s sentiment towards the COVID-19 pandemic. There are missing data for the period from 16 to 23 September 2020 due to the API malfunction issue. The spatial distribution of geotagged tweets, normalised by the total population in each country, is provided in Supplementary Fig. 5. Each tweet record contains a series of attributes (Supplementary Table S9), including 1) *ID* (tweet\_id) as a unique anonymous identifier of a tweet, *date* (date\_time) as the time when a tweet was created, *language* (lang) representing a two-digit language code, *sentiment label* (sentiment\_label) indicating three types of sentiment (positive, neutral and negative), *sentiment confidence* (sentiment\_conf) as the confidence score from 0 to 1 indicating how much confidence a tweet was classified with a given sentiment label, *gender* (gender\_label) indicating the gender of users (details provided in Imran et al.’s publication in 2021 (Imran et al., 2022), as well as five types of country codes. We applied a hierarchical searching algorithm to finalise the location (country) of each tweet among these five types of country codes.

### 2.2. Geotagging

The five types of country codes (Supplementary Table S9) were generated based on five text fields containing toponym mentions,

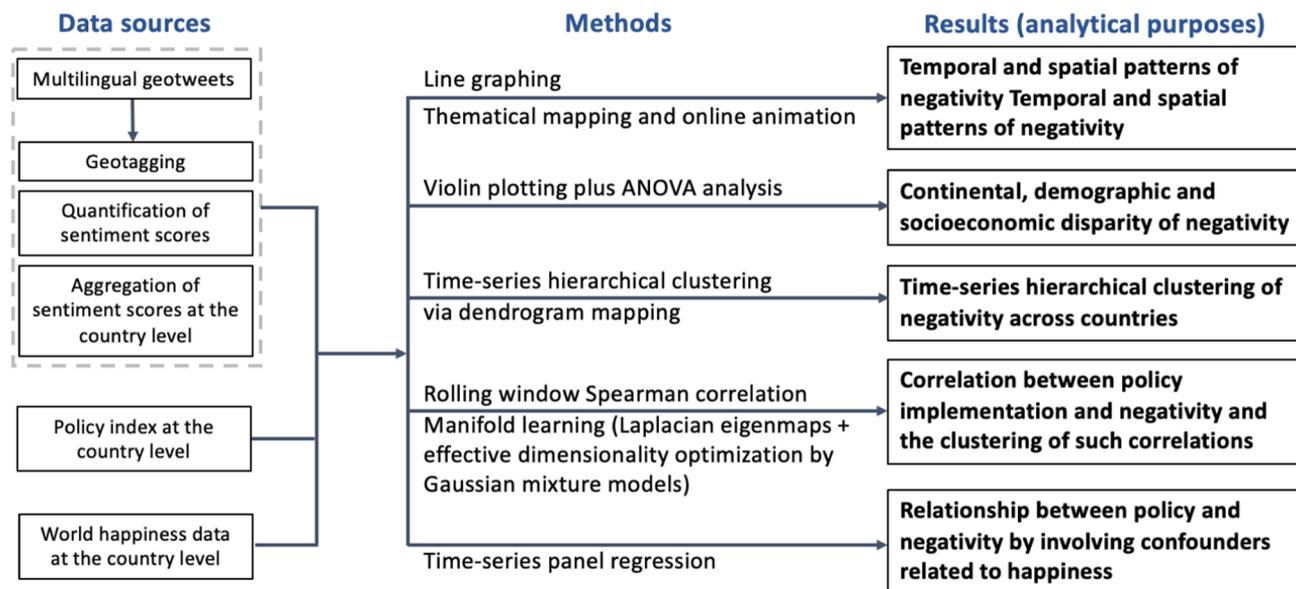


Fig. 1. Research design and analytical workflow.

including geo-coordinates, place tags, user locations, user profile descriptions, and tweet texts. First, the geo-country code was generated based on the geo-coordinates field containing X, Y coordinates as the most accurate location of the user. Geo-coordinates were directly derived from the user's mobile phone device if the user activated the locating function. Second, the place-country code was generated based on the place tags field that were generated by a bounding box of locations that users optionally provide while posting tweets. We then generalized the geo-coordinates and place tags to country names. Third, the user-country code was generated based on the optional field of user locations that allows users to add manually, such as their country, state, and city while posting a tweet. It was then also generalized to a national level. Fourth, the profile-country code was generated based on the user profile description field that carries the user's home country when the user's account was registered. It was saved as a text-based attribute as the potential sources to infer users' locations, on the assumption that users were largely likely to stay in their home countries due to the mobility restriction during COVID-19. Finally, the tweet-country code was detected based on the tweet text field that reflects the actual content of tweets and contain the country that a user discussed about in a tweet (e.g., a user in Italy talking about COVID-19 in China).

In order to finalize the location (country) of each tweet, we applied a hierarchical searching algorithm to assign the first country that was searched in the order of the former four types of country codes (i.e., from geo-country, place-country, user-country to profile-country) as the final country, given that the accuracy of locating users' locations decreased from the geo-country to profile-country. In the end, 1,688,911,319 tweets were excluded without country information (Supplementary Table S10). The tweet-country code was used to indicate tweets differentiated in the country where they were posted and the country which the tweets discussed about. Tweets with such discrepancies were excluded in the later panel regression analysis which only included tweets with the same country where they were posted and discussed about to reflect the impact of local settings (e.g., local policy) on expressed negativity.

### 2.3. Policy index data and mental health indicators

Policy index data was retrieved from the Oxford COVID-19 Government Response Tracker (Hale et al., 2020). It quantified COVID-19 related policies implemented in 180 countries in four dimensions (Supplementary Table S11). Vaccine policies were rarely implemented before March 2021, thus excluded here. Each dimension contains a number of indicators. Four types of policy indices (i.e., the index of government response, containment and health, stringency, and economic support) were constructed based on one or multiple dimensions. In addition, we selected indicators potentially relevant to mental health categorised in the six dimensions outlined in the World Happiness Report (Supplementary Table S12). The yearly World Happiness Report was developed by the United Nations Sustainable Development Solutions Network based on the survey of the state of global happiness, providing a ranking of countries by their happiness levels (Helliwell et al., 2020). It used six key variables to measure the differences in happiness across countries—income, health, social support, generosity, freedom, and trust (Helliwell et al., 2020). Here, we reformulated these six measures of happiness as six dimensions. In each dimension, we selected one or multiple indicators from multiple sources, including WHO (World Health Organization, 2020a), United Nations Population Division (United Nations Population Division, 2020), and the World Happiness Report (Helliwell et al., 2020), which potentially have impacts on the expressed negativity in response to COVID-19. The rationale for selecting these indicators was justified in Supplementary Note 1.

### 2.4. Quantification of sentiment and negativity scores

All tweets in the dataset include sentiment labels indicating positive,

neutral, and negative tweets and sentiment confidence scores ranging from 0 to 1 to indicate the level of confidence a certain tweet was labeled as bearing a positive, neutral, and negative sentiment (Imran et al., 2022). This sentiment confidence score can be used to reflect the intensity of certain types of sentiment. For example, a negative tweet with a sentiment confidence score of 0.9 would have a stronger sense of negativity compared to a negative tweet with a sentiment confidence score of 0.2. After exploring different measures to reflect mental health signals (Supplementary Note 2), we constructed the daily sentiment score at the aggregated country level as Eq. (1), and the daily negativity score as Eq. (2).

$$\text{Dailysentimentscore} = \frac{\sum_0^i t_i^c \times \theta_i^c}{N_c} \quad (1)$$

$$\text{Dailynegativityscore} = (-1) \times \text{Dailysentimentscore} \quad (2)$$

where  $t_i^c$  denotes the sentiment label (positive as 1, neutral as 0, and negative as -1) of tweet  $i$  from a country  $c$ ;  $\theta_i^c$  denotes the sentiment confidence score of  $t_i^c$ ;  $\widehat{\theta}_i^c$  denotes the mean values of  $\theta_i^c$  at the aggregated country level;  $N_c$  is the total number of tweets in a country  $c$ .

### 2.5. Time-series hierarchical clustering and dendrogram mapping

We utilised the dynamic time warping (DTW) approach (Supplementary Note 3) to detect the time-series hierarchical clusters of negativity across countries. The DTW results were visualised as dendrogram maps (Supplementary Note 3). Using X and Y to represent the timeline of negativity of two countries, DTW was calculated as the squared root of the sum of squared distances between each negativity score in series  $X = (x_0, \dots, x)$  and its nearest negativity score in another series  $Y = (y_0, \dots, y)$  (Müller, 2007). The DTW distance from X to Y is formulated as (Müller, 2007):

$$\text{DTW}(x, y) = \min_{\sigma} \sqrt{\sum_{(i,j) \in \sigma} d(x_i, y_j)^2} \quad (3)$$

where  $\sigma = [\sigma_0, \dots, \sigma_k]$  is a list of index pairs of data points (negativity scores) from day 1 to day  $k$  in the country  $j$  and  $i$ ; and  $\sigma$  satisfies the following properties: 1)  $0 \leq i_k < n$  ( $n$  is the number of days with sentiment recorded in the country  $i$ ) and  $0 \leq j_k < m$  ( $m$  is the number of days with sentiment recorded in the country  $j$ ); 2)  $\sigma_0 = (0, 0)$  and  $\sigma_k = (n-1, m-1)$ ; 3) for all  $k > 0$ ,  $\sigma_k = (i_k, j_k)$  is related to  $\sigma_{k-1} = (i_{k-1}, j_{k-1})$ .

### 2.6. Rolling window Spearman's $\rho$

Spearman's  $\rho$  is a nonparametric measure of rank correlation that measures how well the relationship between two series can be described using a monotonic function. Ranging from -1 to 1, a high Spearman's  $\rho$  suggests a similar rank between two series (a value of 1 denotes an identical pattern), while a lower Spearman's  $\rho$  suggests a dissimilar rank (a value of -1 denotes a fully opposed pattern). Here, we calculated the rolling window Spearman's  $\rho$  of negativity levels against the index of policy stringency, government response, and containment health. We discard the economic support index given it contains many missing values in February 2020 and the duration of economic support policy is usually long lasting (same values that last for weeks and even months) that will introduce bias in the correlation analysis. We first processed the missing values using a linear interpolation approach, with 30 days as the maximum number of consecutive missing values allowed to be filled. For missing values that lie outside valid values, we extrapolated them by propagating the closest value. To make sure the statistical robustness of Spearman's  $\rho$ , the window size (i.e.,  $h$ ) was set to 61 (one month prior and one month after the date in the center) to ensure sufficient samples to establish a statistically robust Spearman's  $\rho$  calculation. We obtained window sequences ( $h = 61$ ) of negativity levels, i.e.,  $X = (x_1, \dots, x_h)$  and

sequences of either policy stringency index, government response index, or containment health index, i.e.,  $Y = (y_1, \dots, y_h)$ . We ranked all values in  $X$  and  $Y$  respectively as  $X^r = (x_1^r, \dots, x_h^r)$  and  $Y^r = (y_1^r, \dots, y_h^r)$ . The Spearman rank correlation coefficient between  $X$  and  $Y$  is denoted by  $r_s$  and is calculated by:

$$r_s = \frac{h \sum_{x^r \in X^r, y^r \in Y^r} x_r y_r - \sum_{x^r \in X^r} x_r \sum_{y^r \in Y^r} y_r}{\sqrt{\left( h \sum_{x^r \in X^r} (x^r)^2 - \left( \sum_{x^r \in X^r} x_r \right)^2 \right) \left( h \sum_{y^r \in Y^r} (y^r)^2 - \left( \sum_{y^r \in Y^r} y_r \right)^2 \right)}} \quad (4)$$

where the derived  $r_s$  from each window forms a rolling correlation sequence  $R_s = (r_s^1, \dots, r_s^{n-h+1})$  for each country, which was used as input for subsequent series clustering;  $h$  is the window size set to 61 (one month prior and one month after the date in the center in each period) to ensure sufficient samples to establish a statistically robust Spearman's  $\rho$  calculation. The derived  $r_s$  from each window forms a sequence  $R_s = (r_s^1, \dots, r_s^{n-h+1})$  for each investigated country, serving as input for the subsequent series clustering. Note that Spearman's  $\rho$  returned "NaN" (not a number) if the window sequence in either  $X$  or  $Y$  presented identical values. We further linearly interpolated NaN values in the calculated correlation sequence, with 30 as the maximum number of consecutive NaN values allowed to be filled. Countries were dropped if their correlation sequences still contained NaN values after the interpolation.

## 2.7. Manifold learning by Laplacian eigenmaps

Given the complexity of the correlation between negativity and policies generated by the above rolling window Spearman's  $\rho$ , we employed a manifold learning method for data dimensionality optimization from both the spatial and temporal perspective—which are not easily done by DTW that generates dendrogram maps describing the hierarchical clustering of countries without the spatial and temporal pattern. To cluster country-level  $R_s$ , we first performed manifold learning for  $R_s$  series embedding, aiming to find their low-dimensional representations. The algorithm we employed was Laplacian eigenmaps as a nonlinear manifold learning method used to identify a low-dimensional embedding and to optimally preserve the local structure of a high-dimensional data manifold (Levin et al., 2021), under the assumption that the high-dimensional data resides on a low-dimensional manifold. We first established a data matrix  $\mathbb{R} = [R_s^1, R_s^2, \dots, R_s^m]$  for a total of  $m$  available countries. The values of  $m$  differed in different correlation types:  $m = 147, 159$ , and  $166$  for the correlation of Twitter negativity score against policy stringency index, containment health index, and government response index, respectively. To embed data matrix  $\mathbb{R}$  in a lower  $l$ -dimensional space, the Laplacian Eigenmaps method employs  $l$  eigenvectors of the nearest-neighbors Laplacian graph that corresponds to the smallest non-zero eigenvalues. More details of the specific construction of Laplacian eigenmaps can be found in work by Belkin and Niyogi (Belkin and Niyogi, 2001). In this study, we constructed the affinity matrix by computing a graph of nearest neighbours, with the number of nearest neighbours for graph building set to the integer of  $m/10$ . We selected ARPACK (Lehoucq et al., 1998) as the eigenvalue decomposition strategy. Derived low-dimensional representation of the original data matrix  $\mathbb{R}$  is written as  $\mathbb{R}^{\mathcal{L}} = [{}^L R_s^1, {}^L R_s^2, \dots, {}^L R_s^m]$ .

## 2.8. Effective dimensionality optimization

To obtain an optimized number of effective dimensionality, we employed the trustworthiness metric (Venna and Kaski, 2001) that quantifies to what extent the local structure of embedded space, i.e.,  $\mathbb{R}^{\mathcal{L}}$ , was retained from the original higher-dimensional space, i.e.,  $\mathbb{R}$ , after the dimensional reduction via Laplacian Eigenmaps. In general, trustworthiness penalized unexpected nearest neighbors in the embedded space in proportion to their rank in the original high-dimensional space.

The calculation of trustworthiness follows:

$$T = 1 - \frac{2}{mq(2m-3q-1)} \sum_{i=1}^m \sum_{j \in \mathcal{N}_i^q} \max(0, (r(i,j) - q)) \quad (5)$$

where for a total of  $m$  samples,  $\mathcal{N}_i^q$  denotes the  $q$  nearest neighbors of sample  $i$  in the  $l$ -dimensional embedded space, and sample  $j$  corresponds to its  $r(i,j)$ -th nearest neighbor in the original high-dimensional space. Following the number of nearest neighbors in the Laplacian graph building, we set  $q$  to integer of  $m/10$ . We calculated trustworthiness  $T$  as a function of Laplacian Eigenmap embedding dimensionality (i.e.,  $l$ ). We assumed the optimal  $l$  setting lied in a position where the relative costs to increase  $l$  are no longer worth the corresponding boosts in  $T$ . Thus, the optimal setting of  $l$  can be determined via a knee-point detection algorithm on a series of  $T$  values that correspond to different  $l$  settings. We implemented the knee-point detection using the *Kneedle* algorithm from the *kneed* package (Satopaa et al., 2011). We smoothed the  $T$  series via polynomial fitting and set the curve type to "concave" and direction to "increase". The optimal embedding dimensionality for the window Spearman correlations of Twitter negativity score against the index of policy stringency, government response, and containment health are 18, 20, and 19, respectively (Supplementary Figure S3).

## 2.9. Gaussian mixture model (GMM) and cluster number optimization

Following the generation of low-dimensional structure  $\mathbb{R}^{\mathcal{L}}$  optimized by trustworthiness metric and knee-point detection, we used Gaussian Mixture Model (GMM) to perform clustering on  $\mathbb{R}^{\mathcal{L}}$ . GMM is a probabilistic model on the assumption that all the data points are generated from a mixture of Gaussian distributions with parameters governing the cluster centroid and covariance structure. The GMM model can be written as  $P(\mathcal{L} R_s^i | \theta) = \sum_{k=1}^K \pi_k N(\mathcal{L} R_s^i | \mu_k, \Sigma_k)$ , where  $\theta$  stands for model parameters,  ${}^L R_s^i$  denotes a certain embedded low-dimensional representation of country  $i$  after Laplacian eigenmaps embedding,  $\pi$  denotes the probability mass function of  $z_i$ , whose posterior probability  $P(z_i = k | {}^L R_s^i, \theta)$  that point  ${}^L R_s^i$  belongs to cluster  $k$  can be computed as:

$$\begin{aligned} P(z_i = k | {}^L R_s^i, \theta) &= \frac{P(z_i = k | \theta) P({}^L R_s^i | z_i = k, \theta)}{\sum_{k'=1}^K P(z_i = k' | \theta) P({}^L R_s^i | z_i = k', \theta)} \\ &= \frac{\pi_k \mathcal{N}({}^L R_s^i | \mu_k, \Sigma_k)}{\sum_{k'=1}^K \pi_{k'} \mathcal{N}({}^L R_s^i | \mu_{k'}, \Sigma_{k'})} \end{aligned} \quad (6)$$

Following the method by Levin et al. (2021), we further identified the optimized number of GMM components (i.e., clusters) using the Bayesian information criterion (BIC) based on a penalized form of the log-likelihood (more heavily on model complexity compared with AIC). That is to say, with the involvement of additional components, the penalty term for the number of estimated parameters was subtracted from the log-likelihood:  $BIC = -2 \times \ln(Q) + p \ln(m)$ , where  $Q$  denotes the maximum value of the likelihood function for the model,  $p$  denotes the number of estimated parameters in the model, and  $m$  denotes the total number of data points (i.e., the  $m$  available countries). We selected the cluster number with a minimized BIC value. The optimal number of GMM components (i.e., clusters) for the embedded space of the correlation between Twitter negativity score against the index of policy stringency, government response, and containment health are 3, 3, and 4, respectively (Supplementary Figure S4).

## 2.10. Time-series panel regression

After the original tweet data at the individual level was aggregated by day and by country and merged with country-specific mental health indicators (Supplementary Table S14), we forged a panel dataset containing time-variant variables (e.g., negativity scores and policy indices)

and time-invariant variables (country-specific mental health indicators). Variables that are potentially relevant to the level of negativity were categorised in the six dimensions outlined in the World Happiness Report (Supplementary Note 4). In addition, it is likely that the presence of complex behavioural, contextual and epidemiological variables may also affect the viral transmission in the population, which may further influence the level of negativity. These latent variables were difficult to be fully captured and quantified. As such, we employed a time-series panel regression to estimate the relationship between the level of negativity and the selected variables, given that the panel regression is flexible to account for time-invariant country attributes and time-variant confounders to reduce the modelling bias (Brüderl and Ludwig, 2015). In order to ease the concern of potential multicollinearity amongst the variables, we first conducted a pairwise correlation of variables (Supplementary Figure S7) and then constructed the panel regressions as below (Torres-Reyna, 2007):

$$Y_{c,t} = \mu_i X_{c,t,i} + \sigma_j X_{c,j} + \alpha_c + \beta_i + \varepsilon_{c,t} \tag{7}$$

where  $Y_{c,t}$  is the level of negativity in country  $c$  on date  $t$ ;  $X_{c,t,i}$  is within-country time-variant variables (i.e., policy stringency index) measured at the date which  $Y$  is observed and  $\mu_i$  ( $i = 5$ ) are the coefficients for each  $X_{c,t,i}$ ;  $X_{c,j}$  is the country-specific time-invariant variables selected from the World Happiness Report (i.e., GDP per capita, life expectancy, medium age, and population density) and  $\sigma_j$  are the coefficients for each  $X_{c,j}$ ;  $\alpha_c$  denotes country-specific effects while  $\beta_i$  denotes date-specific effects at the global level or country-specific time trends;  $\varepsilon_{c,t}$  are standard errors at the country level.

2.11. Hausman test

Another concern stemming from unobservable confounders is the uncertainty that if country-specific and time-specific effects that may capture some of the unobserved variations were fixed or random (Torres-Reyna, 2007). Thus, we tested out both fixed and random effects and selected better models via the comparison of the modelling performances. The fixed effects model is based on the primary assumption that there are certain attributes of a given individual country that do not vary over time. Such attributes might or might not be correlated with the individual dependent variable (the level of negativity). While in fixed effects model, we have controlled for differences between individual countries. Another important assumption of the fixed effect model is that country-specific time-invariant variables are unique to the individual and should not be correlated with other time-variant variables. However, we were less sure about if such variables were constant across countries or changed over time. Thus, we also constructed a random effects model considering these country-specific variations as well as time dependent variations. The Hausman test, as shown as below,

was used to test out which models performed better:

$$H = (\hat{b}_r - \hat{b}_f)' [Var(\hat{b}_r) - Var(\hat{b}_f)]^{-1} (\hat{b}_r - \hat{b}_f) \tag{8}$$

where  $b_r$  is a vector of coefficients generated by the random effects model while  $b_f$  is a vector of coefficients generated by the fixed effects model. If the Hausman test result is with  $p < 0.01$  then the fixed effects model performs better than the random effects model; otherwise the random effects model is more suitable.

We tested both fixed and random effects panel regression models and selected the fixed effects model due to its superior modelling performance by Hausman Test (Supplementary Note 5) (Torres-Reyna, 2007). We ran the fixed effects panel regression, first taking each policy index as the single independent variable, and then ran the model again by adding a set of country-specific mental health indicators (Supplementary Figure S7). The later model had an improved model performance with the increase of  $R^2$  from 0.184 (Supplementary Table S15) to 0.37 (Table 1). It reflects that the additional involvement of country-specific mental health indicators contributed to better explaining the variation of negativity.

3. Results

3.1. Temporal and spatial patterns of negativity towards COVID-19

We commenced with exploring the temporal change of negativity towards COVID-19 based on the daily negativity scores smoothed in a 7-day window (detailed in the Methods) by gender (Fig. 2A) and country (Fig. 2B). In general, the daily average of negativity scores (green dash line in Fig. 2A) increased in February but decreased from March to May 2020, and remained minor fluctuations afterwards till March 2021. The negativity of female Twitter users tended to be lower than that of male users from the middle of March to early May 2020 but displayed a minor difference with that of males after May 2020. For the top 10 countries with the largest cumulative number of COVID-19 cases by 31 March 2021 (Supplementary Table S1), France (yellow line in Fig. 2B) displayed a higher level of negativity over all other countries from late February to the mid of March 2020; from April to the end of December 2020, the U.S. and Italy presented relatively higher levels of negativity among the ten countries. All other countries (e.g., Brazil, France, Russia, U.K., Spain, and Germany in Fig. 2B) ranged in the middle level of negativity with minor differences between each other. Turkey and India had relatively lower levels of negativity, especially from November 2020 to the end of March 2021.

The spatial patterns of negativity have been visualised as a user-interactive animation in the project website (indicated in ‘Data Availability’) based on the daily country-level tweet data ( $N = 90,489$  as 217

**Table 1**  
Fixed effects panel regression for four policy indices over the whole research period.

	$\beta$	95 % CI						
Stringency index	-0.052***	(-0.058, -0.047)						
Economic support index			-0.029***	(-0.035, -0.024)				
Government response index					-0.070***	(-0.077, -0.064)		
Containment and health index							-0.060***	(-0.066, -0.053)
GDP per capita	0.720***	(0.643, 0.796)	0.757***	(0.681, 0.833)	0.709***	(0.633, 0.786)	0.719***	(0.643, 0.796)
Population density	2.626***	(2.34, 2.911)	2.884**	(2.599, 3.168)	2.723**	(2.439, 3.008)	2.742**	(2.457, 3.026)
Life expectancy	-1.637***	(-1.736, -1.538)	-1.660***	(-1.759, -1.561)	-1.627***	(-1.725, -1.528)	-1.640***	(-1.738, -1.541)
Medium age	-4.655**	(-5.2, -4.109)	-5.199**	(-5.741, -4.657)	-4.875**	(-5.418, -4.332)	-4.910**	(-5.453, -4.367)
Age 65 and above	6.750***	(6.279, 7.22)	7.314***	(6.848, 7.78)	6.907***	(6.44, 7.375)	6.950***	(6.482, 7.418)
Hospital beds	-0.267**	(-0.35, -0.184)	-0.348***	(-0.431, -0.266)	-0.294***	(-0.377, -0.212)	-0.293***	(-0.376, -0.21)
Generosity	1.119**	(1.069, 1.168)	1.191**	(1.143, 1.24)	1.118**	(1.068, 1.167)	1.132**	(1.082, 1.181)
R-square	0.378		0.377		0.378		0.378	
F-statistic	221.769		221.020		222.005		221.819	
Number	70,964		70,964		70,964		70,964	

Note: \*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$ ;  $\beta$ : standardised coefficient; CI: confident interval.

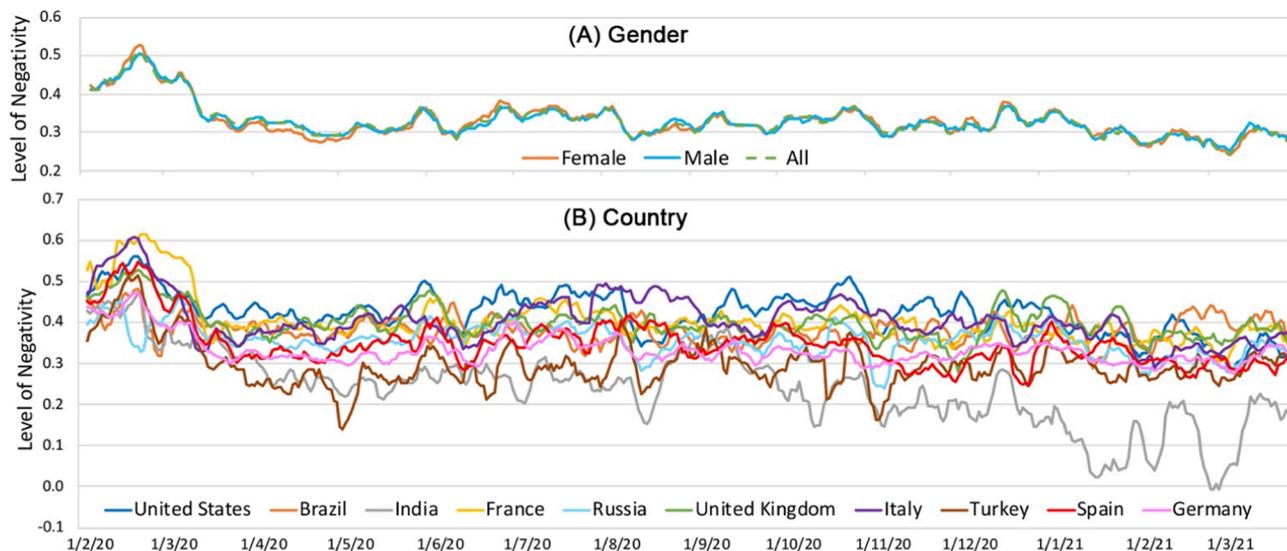


Fig. 2. Temporal change of negativity towards COVID-19 by (A) gender; (B) top 10 countries with the largest cumulative number of COVID-19 cases by 31 March 2021 (Supplementary Table S1).

countries by 417 days). We selected six static patterns of negativity on six particular dates (Fig. 3). On 1 February 2020, a high level of negativity was observed in Iran, Australia, and China, where the first COVID-

19 case was reported, and the early outbreak of COVID-19 appeared. Some small countries (e.g., island countries) reported having a high level of negativity, including Samoa, Liechtenstein, Tonga, and the

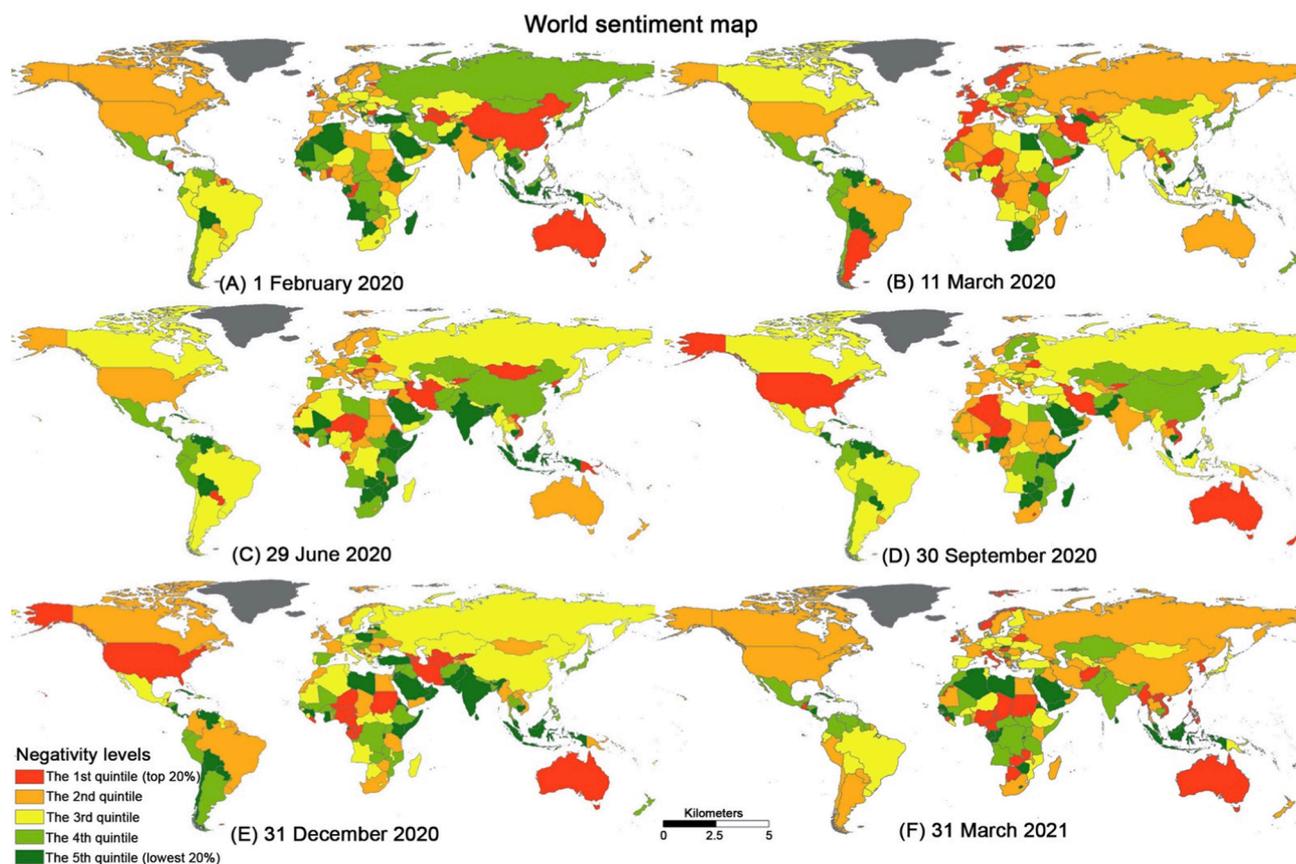


Fig. 3. Spatial pattern of negativity on six selected dates: (A) 1 February 2020 as the first day of the research period to represent the negativity in the pre-pandemic phase; (B) 11 March 2020 as the date on which the World Health Organization (WHO) officially declared the global COVID-19 pandemic; (C) 29 June 2020 as the six-month anniversary of the COVID-19 outbreak marked by WHO; (D) 30 September 2020 as the date by which the COVID-19 curve flattened substantially; (E) 31 December 2020 as the date when the Pfizer / BioNTech vaccine was firstly received the validation from WHO; (F) 31 March 2021 as the final date of the research period. The levels of negativity were classified to quintiles presented in an orange-green colour ramp; the high quintile with the orange colour indicated high levels of negativity; reversely dark green indicated the low levels of negativity. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Republic of the Congo, which may be difficult to be observed on the global map (Supplementary Table S2). On 11 March 2020, high levels of negativity were observed in Iran, Argentina, some European (e.g., U.K., France, Italy, Spain, Norway, and Sweden) and African countries (e.g., Niger, Kenya, and Yemen). On 30 September, 31 December 2020, and 31 March 2021, higher levels of negativity were observed in the US, Australia, Iran, Afghanistan, and some African countries (e.g., Nigeria and Sudan). Overall, the spatial distribution of countries with high levels of negativity was dispersal at the global scale, although a number of large countries (e.g., the U.S. and Australia) displayed high levels of negativity consistently.

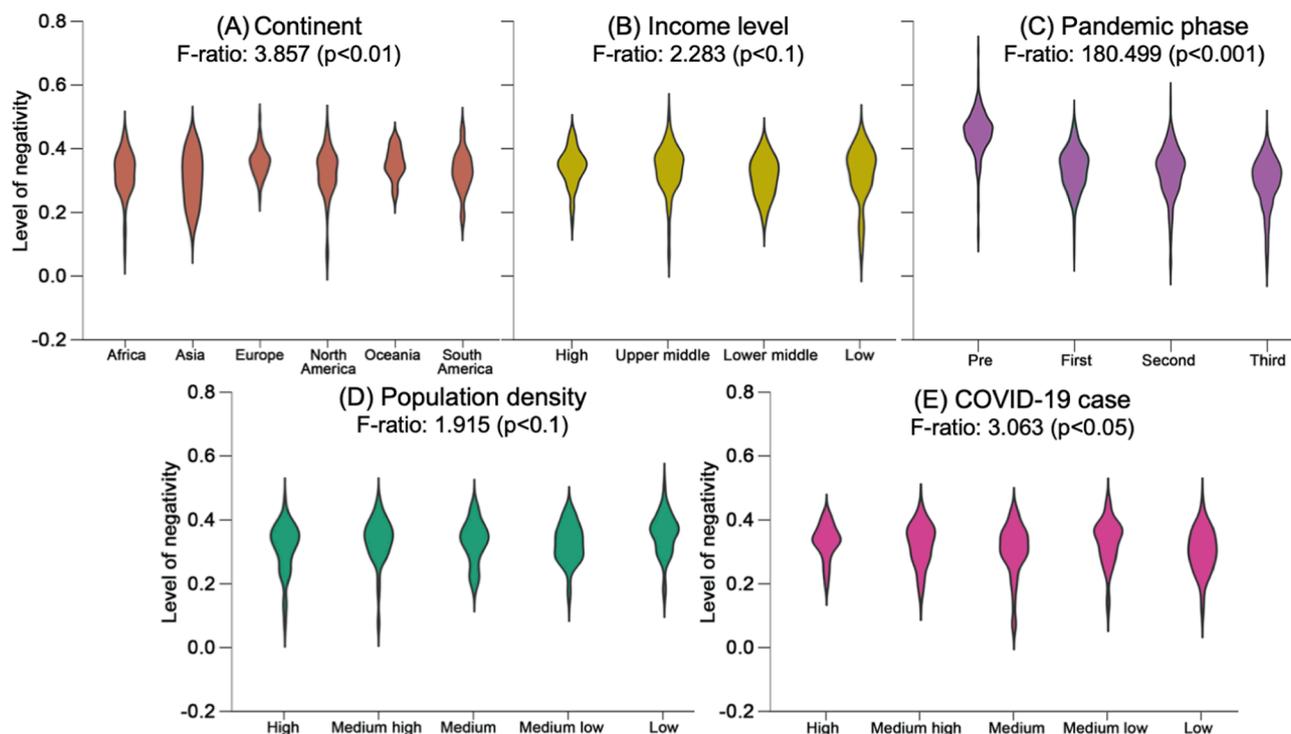
### 3.2. Continental, demographic and socioeconomic disparity of negativity

We conducted a set of one-way ANOVA analyses via cross-tabulation and violin plotting (see Methods) to examine the disparity of negativity across different groups in Fig. 4 (statistical details in Supplementary Table S3). Grouped by continent (Fig. 4A), the mean of negativity scores in Europe was the highest (0.357 in Supplementary Table S3), followed by Oceania (0.354) and South America (0.328); while the range of negativity scores in North America (0.071–0.452) was wider than that in other continents, indicating a larger fluctuation of negativity observed in North America. Grouped by income levels (Fig. 4B, definition

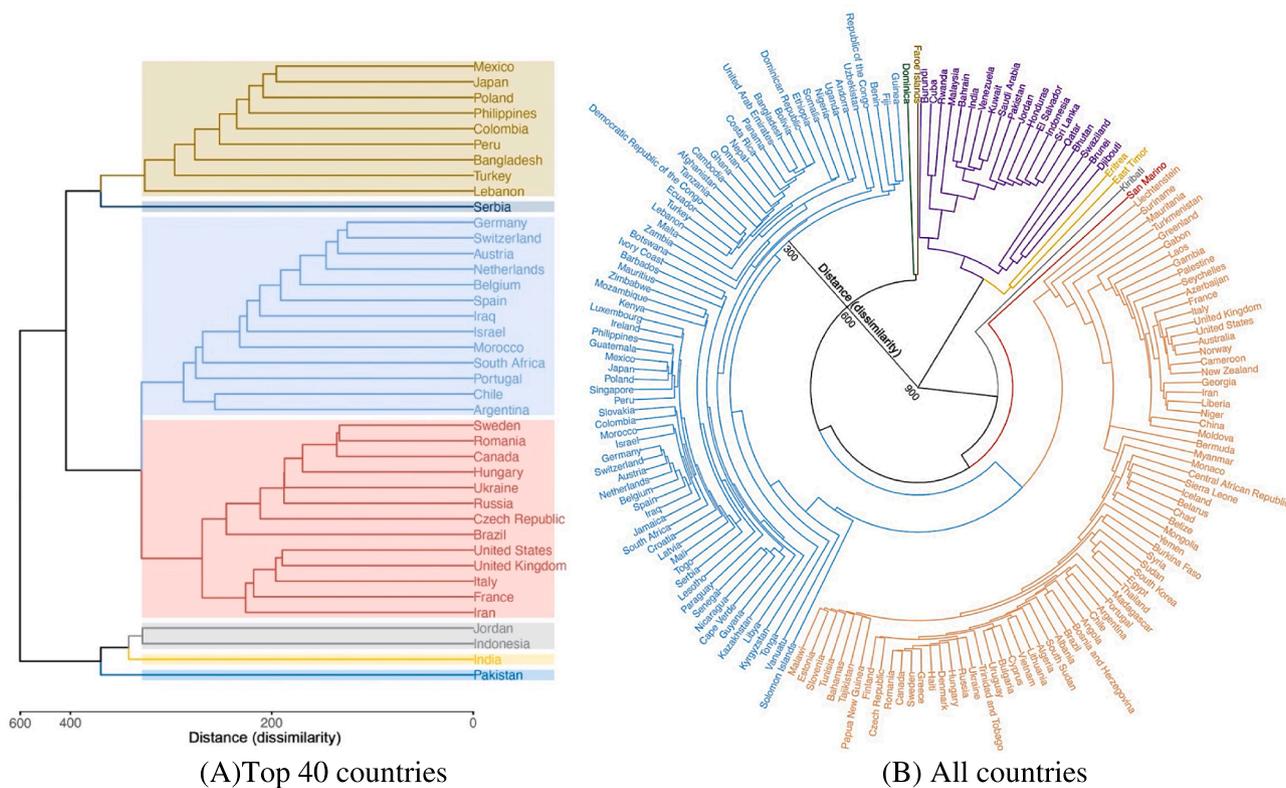
provided in the Method), the mean values of negativity in high income (0.341) and upper-middle-income countries (0.336) were higher than that in low-income (0.322) and lower-middle-income countries (0.311); while the range of negativity scores in low-income countries (0.078–0.445) was wider than that in other countries. Grouped by pandemic phases (Fig. 4C), it was obvious to see that the mean value of negativity scores decreased gradually from 0.449 in the pre-pandemic period (February 2020) to 0.294 in the third phase of the pandemic (January to March 2021). It reflected that the worldwide sentiment becomes more positive along the pandemic timeline. Grouped by population density (Fig. 4D), the low-density countries (defined in the Method) had a higher level of negativity (0.352), followed by medium-high (0.332) and medium-low density countries (0.325). Grouped by COVID-19 cases (Fig. 4E), the ranges of negativity in the medium (0.071–0.425) and low level (0.169–0.497) of viral infection were wider than ranges in other countries. In summary, there existed geographic, demographic, socioeconomic, and phasic disparities of negativity among the 217 countries.

### 3.3. Time-series hierarchical clustering of negativity

We further examined the time-series clustering of negativity across the top 40 countries with the largest accumulative number of COVID-19



**Fig. 4.** Disparity of negativity towards COVID-19 across five types of groups: (A) by continents, including six continents (excluding Antarctica)—Asia, Africa, Europe, Oceania, North and South America; (B) by country income levels, defined by the World Bank that classified the world's economies to four income groups based on the gross national income per capita in USD on July 1 of each year (i.e., 2020 in this case)—low income (<USD 1,035), lower-middle income (USD 1,035–4,045), upper-middle income (USD 4,046–12,535), and high income (>USD 12,535); (C) by pandemic phases—the pandemic timeline from February 2020 to March 2021 was delineated as four phases: February 2020 as the pre-pandemic period (named as 'pre') given that WHO declared COVID-19 as the global pandemic on March 11, 2020 (World Health Organization, 2020b); the first period of the pandemic (named as 'first') from March 2020 to June 2020 during which the COVID-19 rapidly spread out to 200 countries (<https://covid19.who.int/>); the second period of the pandemic (named as 'second') from July to December 2020 during which the spread of COVID-19 slowed down; the third period of the pandemic (named as 'third') from January to March 2021 during which the COVID-19 vaccines started rollout after the Pfizer/BioNTech vaccine received the validation from WHO on 31 December 2020 (World Health Organization, 2020c); (D) by population densities, based on the number of populations per square km—the top quintile of 80–100 % as the high level, 60–80 % as the medium high level, 40–60 % as the medium level, 20–40 % as the medium low level, and 0–20 % as the low level; (E) by COVID-19 cases, based on the accumulative number of COVID-19 confirmed cases by 31 March 2021—the top quintile of 80–100 % as the high level, 60–80 % as the medium high level, 40–60 % as the medium level, 20–40 % as the medium low level, and 0–20 % as the low level. The statistical summary of each violin plot was provided in Supplementary Table S3. The horizontal width of each violin plot represented the density of countries at a certain level of negativity; the vertical range of each violin plot indicated the range of negativity scores at the country level. F-ratios, as the ratio of two mean square values, ranged from 1.915 in Fig. 3D to 180.499 in Fig. 3C. F-ratios were larger than 1 with at least p-value < 0.1, reflecting the variations of negativity across different groups.



**Fig. 5.** Time-series hierarchical clustering dendrogram for **(A)** the top 40 countries with the largest accumulative number of COVID-19 cases by 31 March 2021; **(B)** all countries. The clustered countries were nested and organised as a tree (or circle) as a meaningful classification scheme. Each node in the cluster tree (or circle) contained a group of similar data. In **(A)**, clusters at one level are joined with clusters at the next level up, using a degree of dissimilarity (X-axis). Each joint of two clusters was reflected on the graph by the splitting of one horizontal line into two. The short vertical bar represents the horizontal position of the split, indicating the distance (dissimilarity) between two clusters. In **(B)**, it was visualised in a circular form due to a large number of countries that were not well presented in a vertical form as **(A)**. The number of countries in **(B)** was 176. There were 41 out of the original 217 countries excluded here, given they were largely small countries (e.g., island countries or colonial territories) with missing values in February 2020.

cases by 31 March 2021 (Fig. 5A) and the negativity levels among all the 217 countries (Fig. 5B). There were seven clusters detected in Fig. 5A based on the dissimilarity of the time-variant patterns of negativity along the pandemic timeline. The bottom group, including Jordan, Indonesia, India, and Pakistan, had similar patterns with a lower level of negativity, while the two large groups of countries (light blue and red colour) in the middle had similar patterns with a higher level of negativity (Fig. 5A). Among all the 217 countries (Fig. 5B), there were eight clusters detected. The purple group (from Burundi to Djibouti) had the lowest level of negativity, meanwhile keeping similar patterns with the yellow group (Eritrea and East Timor). There were two large groups of countries (85 % of the total 176 countries marked in dark blue and orange colours) with similar patterns of negativity; the orange group (from Liechtenstein to Malawi) had a relatively higher level of negativity than the dark blue group (from Solomon islands to Guinea).

### 3.4. Relationship between policy implementation and negativity

We observed some common findings on the correlations between policy implementations and negativity across three policy indices (Fig. 6 and Supplementary Fig S1 and S2). The correlations between all three policy indices and negativity reached the bottom (Fig. 5e and Supplementary Fig. S1d and S2d) and were strongly negative in mid-March 2020 when the WHO declared COVID-19 as a global pandemic and most countries started to aggressively strengthen their policy implementation. That is, the worldwide strengthened policy interventions in mid-March 2020 led to reduced negativity towards COVID-19. With the stabilizing and continuity of policies, the negative correlations between policy implementation and negativity gradually weakened and started

to fluctuate around 0 since May 2020. Despite the similarity in the general trend around mid-March 2020, country-wise disparities were notable from the clustering results. For example, Australia, the UK, and Egypt, among other 21 counties of Cluster 3, presented different correlations between negativity and the government response index, compared to Clusters 1 and 2 (Fig. 6), evidenced by their over-bounced correlations in May (from  $-0.75$  in mid-March to  $0.2$  in May) and consistent, weak positive correlations (around  $0.15$ ) from late 2020 to early 2021. As for the containment health index (Supplementary Fig S2), countries in Cluster 4 (Indonesia, Sweden, South Africa, among other 12 members) were featured by their positive correlations with coefficients of  $0.3$  from mid-July to mid-August 2020 and low to  $-0.2$  from January to February 2021.

Differing from the above Laplacian eigenmap embedding and clustering analysis, we further examined the relationship between negativity and policy implementation by involving country-specific mental health indicators reformulated from the World Happiness Report (see Method). We found that all four policy indices were negatively ( $p < 0.001$ ) associated with the level of negativity in the whole research period (Table 1). This means that countries with stronger government response and more efficient policy implementation (in terms of economic support, containment and closure, and health intervention) to cope with COVID-19 tended to have a lower level of negativity. All country-specific mental health indicators reformulated from the World Happiness Report were significantly (at least  $p < 0.01$ ) associated with the level of negativity. More specifically, the countries with higher gross domestic product (GDP) per capita ( $\beta = 0.709-0.757$ ), higher population density ( $\beta = 2.626-2.884$ ), and larger proportions of elderly populations ( $\beta = 6.75-7.314$ ) tended to have higher levels of negativity. It may be

Twitter Negativity and Government Response Index

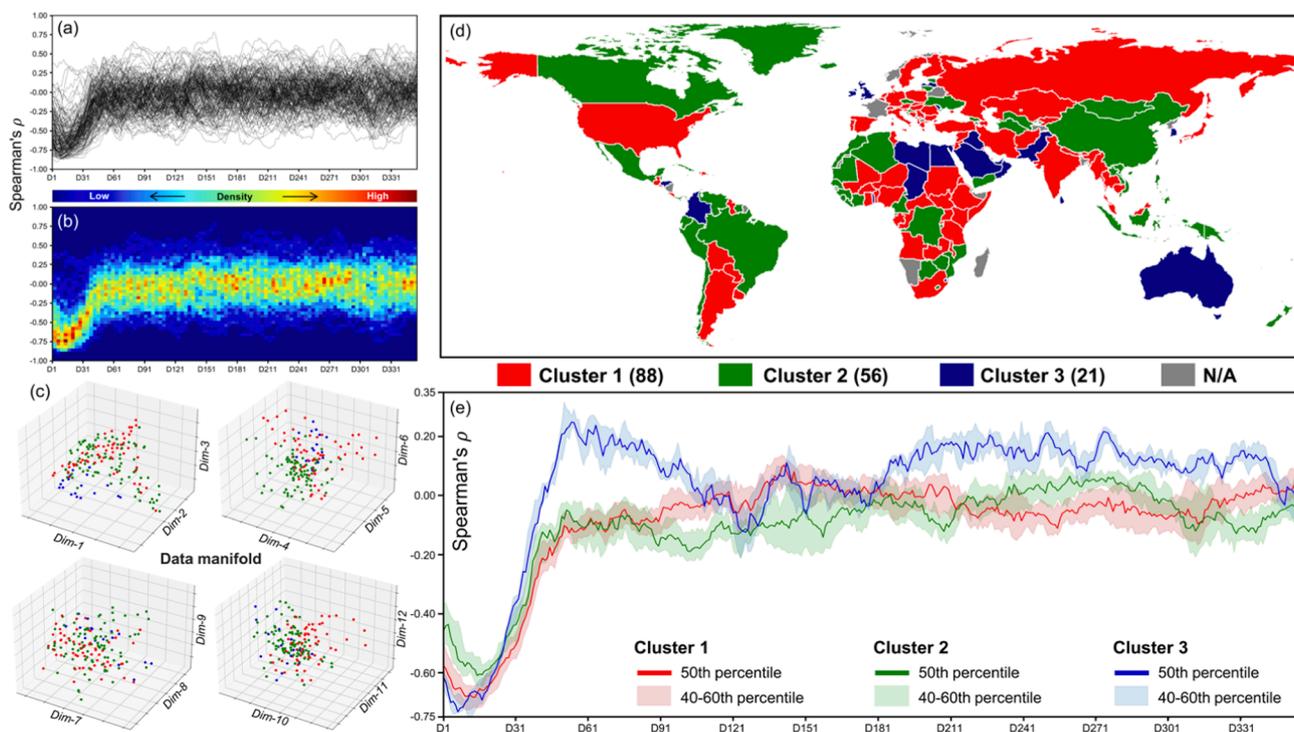


Fig. 6. Country-wise patterns of the correlation between government response index and negativity. (a) the raw sequences of rolling window Spearman's rank correlation with a window size of 61 days (see Methods). Ranging from  $-1$  to  $1$ , a high value of correlation coefficients suggested a similar rank, with the value of  $1$  denoting an identical pattern, while a lower value suggested a dissimilar rank, with the value of  $-1$  denoting a fully opposed pattern. A total of 165 countries were involved in the analysis. (b) the heat map of the raw sequences of rolling window Spearman's rank correlation. High and low concentrations were marked as red and blue, respectively. (c) The Laplacian eigenmaps algorithm was applied to embed the original data matrix to a low-dimensional space with an optimal dimensionality of 20 (Supplementary Figure S3) in a nonlinear manner (See Method). The colors corresponded to an optimized three-component GMM model applied to the 18-dimensional embedding space (Supplementary Figure S4). The first 12 dimensions were visualized with dimensional groups of (1, 2, 3), (4, 5, 6), (7, 8, 9), and (10, 11, 12). (d) The spatial distribution of three identified clusters. (e) The identified three time-series clusters with their median value sequences (i.e., 50th percentile) represented using solid lines of red (Cluster 1), green (Cluster 2), and blue (Cluster 3). The shaded areas represented the corresponding 40-60th percentile uncertainty. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

explained by the fact that most developed countries (e.g., U.S., U.K., and European countries) have more negative tweets posted against COVID-19, and such negative emotions are less obvious in small countries which may have fewer Twitter users. It could also be possible that people in developed countries do not often face negative lifestyle changes (e.g., mobility and travel restrictions) and thus may be more sensitive to these restrictions, which were reported to be associated with high

psychological distress (Alon-Tirosh et al., 2021). It was in contrast with what was reported by the World Happiness Report that wealthier countries tended to be happier and more positive (John F. Helliwell, 2021). In other words, the public's mental signals in response to a particular phenomenon or circumstance may function differently compared to long-term evaluations of mental status. Furthermore, countries with longer life expectancy ( $\beta = -1.66$ – $-1.627$ ) and stronger

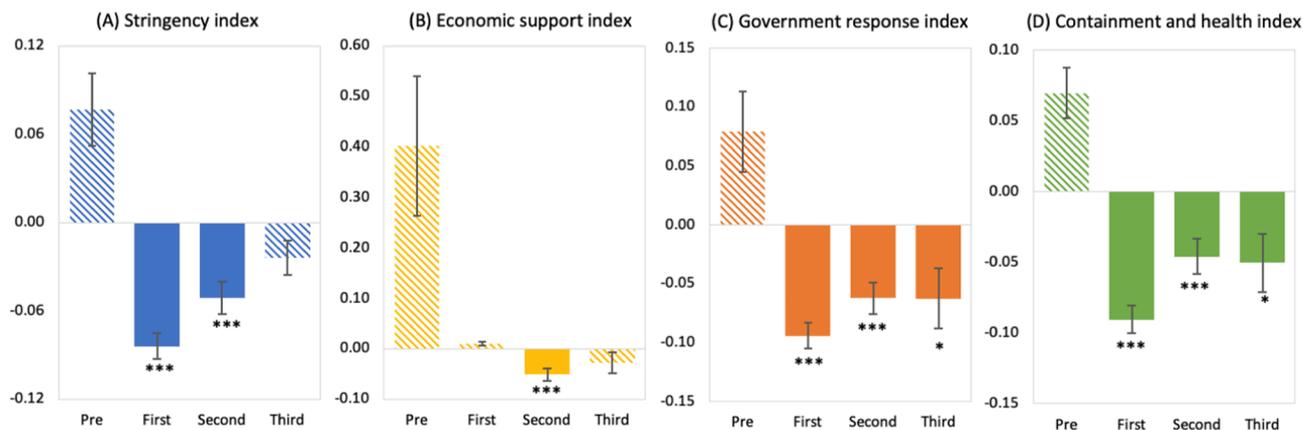


Fig. 7. Coefficients of four policy indices and COVID-19 variables across four pandemic phases. Dash bars indicated the insignificant coefficients; solid bars indicated significant coefficients at various levels:  $**p < 0.001$ ;  $*p < 0.01$ ;  $p < 0.05$ ; X-axis indicated the four phases of the pandemic—pre-pandemic (February 2020), the first phase (March to June 2020), the second phase (July to December 2020), and the third phase (January to March 2021). Y-axis indicated the magnitude of coefficients generated by the fixed effects panel regression (Supplementary Table S4–S7). Error bars represented the 95 % confidence intervals.

medical capacity ( $\beta = -0.348$ – $-0.267$ ) tended to have lower levels of negativity. It was within our expectation that the availability of medical resources and capacities (e.g., hospital beds, hygiene facilities, and clinicians) would reduce the public's fear, concern, and pessimism about the pandemic.

In addition, we ran a set of panel regression models across four pandemic phases to reveal the dynamics between policy implication and negativity with the fixed effects of country-specific mental health indicators (Supplementary Table S4 to S7). The coefficients of four policy indices (Fig. 7) in the pre-pandemic phase (February 2020) were insignificantly ( $p$  greater than 0.1) associated with the level of negativity, but three of them (except for the economic support index) became significantly ( $p < 0.001$ ) associated with the level of negativity in the first and second phase of the pandemic (from March to December 2020). In the third phase (January to March 2021) of the pandemic, only government response, containment, and health indexes were significantly ( $p < 0.05$ ) associated with the level of negativity, indicating that negativity evolved as the function of government response and containment and health interventions.

## 4. Discussion

### 4.1. Principal results

Our study contributes the global investigation of mental health signals to COVID-19 across multiple languages over the pandemic timeline. We found that negativity shifted over time and varied across geographic, demographic, and socioeconomic contexts. Furthermore, we found temporal clusters of negativity across countries and along the epidemic timeline. In addition, we identified that changes in government policy were associated with shifts in negativity though this association weakened in the later stages of the pandemic. More specifically, the first finding was that the mental health signals to COVID-19 expressed on social media were most negative in February 2020, following which the negativity gradually disappointed between March and June 2020 and then remained relatively stable post-March 2021. This temporal shift of mental health signals has been observed in other studies using social media data in single countries (Hussain et al., 2021; Jang et al., 2021; Wang et al., 2022). Second, there were important geographic, demographic, and socioeconomic disparities of negativity across continents, different levels of a nation's income, population density, and the level of COVID-19 infection. This was a unique finding that has not been revealed in the current literature, although some existing studies presented the geographic and socioeconomic disparities of sentiment towards the pandemic within one country (Hu et al., 2021; Rahman et al., 2021; Wang et al., 2022). Third, we detected eight clusters of countries in terms of the time-variant pattern of negativity. The two largest groups containing 85 % of all countries displayed similar time clustering patterns, while the remaining countries presented clear discrepancies across cluster types. We did not observe the similarity of negativity patterns among countries that were geographically adjacent to one another; instead, these countries with similar negativity patterns in a single cluster type tended to be geographically randomly distributed. This was a unique finding that has not been revealed in the current literature.

### 4.2. Comparison with prior work

We found a significant negative association between four types of policy indices and negativity that was observed in the first two months of the pandemic (March and April 2020). This goes some way to highlight that countries that had stronger government response and better implementation of economic support, containment and closure, and health policies to cope with COVID-19 tended to have lower levels of negativity. A number of existing studies echo this finding in that positive sentiments were reported to be related to quarantine, social distancing,

stay-at-home policies, and masks mandates, especially during the early stage of the pandemic (Cheng et al., 2021; Jang et al., 2021; Saleh et al., 2021; Wang et al., 2020). Additionally, the negativity towards COVID-19 dropped and shifted after the early stage of the pandemic, which indicated the public's broad acceptance of the pandemic allied with the broad acceptance of government policies enforced to contain and manage its impact (Naseem et al., 2021). This finding also partially ran counter to some studies concluding that social restriction policies were associated with increased mental health issues (e.g., depression, mental disorders, and feeling of loneliness and isolation) (Díaz and Henríquez, 2021; Suratnoaji et al., 2020). However, these studies gave relatively more attention to government containment and closure policies (e.g., staying-at-home orders, gathering restrictions, and school/workplace closure), rather than the economic (e.g., debt relief for households, income support) and health (e.g., contact tracing, vaccination policy) support-related policies which may function differently and offset the adverse impact of social restriction policy on mental health (Cheng et al., 2021; Jang et al., 2021; Saleh et al., 2021; Wang et al., 2020). It may be also explained by the diversity of culture and religions across countries which have been observed to influence perceived stress to the COVID-19 pandemic (Ting et al., 2021). Moreover, the effect of policy implementation on mental health became marginal at the later stage of the pandemic.

### 4.3. Policy impletations

Based on our findings, we propose several policy implications to monitor, control, and balance the dynamics between public mental health and policy implementation in and beyond the pandemic period. First, international health authorities can employ our framework, together with the support of our tweet sentiment data and methods sharable to the public to track the public's mental health and to design strategies on mental health in the response to future public health emergencies. Second, the implementation of health intervention policies and the provision of mental health services should be adjusted efficiently in accordance with the public's real-time mental health status during different pandemic phases. Third, the evaluations of real-time (e.g., via Twitter) and survey-based mental health status (e.g., via the World Happiness Report) are both needed. Supplementary to the survey-based evaluations, monitoring real-time mental health status on social media, usually at a large spatial and temporal scale, captures the instant reaction of the public to certain circumstances. It enhances the design and implementation of policies in response to mental health issues during public health emergencies and lowers the cost of mental health evaluations, which are expensive and time-consuming via surveys. Fourth, digital platforms and social media channels need to be incorporated to disseminate accurate information and guidance to improve mental health efficiently and effectively.

### 4.4. Limitations

There are a number of study limitations that can be addressed in future studies to extend our findings. First, Twitter users can not represent the overall population as acknowledged by many studies (Blank, 2017; Tufekci, 2014). For example, we may miss the information from the elderly and people who do not use digital devices and social media in this study. There also exists regional differences in Twitter users to make our analysis less representative in some countries (e.g., China). Future studies need to incorporate different social media platforms (e.g., Weibo used in China) to increase the data representativeness and coverage. Second, the quantification of negativity scores may be subject to systemic bias caused by the nature of qualitative content data and the usage of multiple languages. There may be discrepancies between the genuine thoughts the actual expressions of Twitter users given that the routine of social expressions may be different across languages. Third, there may be time-lag effects of policy implementation on mental

health, which need to be considered in future studies. Finally, given mental health is multifactorial, future work could consider incorporating explanatory variables in psychological and social levels in analytical models (e.g., culture and personality) to better understand its causes (Ahmed, 2010).

## 5. Conclusions

To conclude, our study makes unique contributions to the literature as below. First, to our best knowledge, we provide the first global and multilingual evaluation of the public's real-time mental health signals to COVID-19 at a large spatial and temporal scale. The data and the methods of data generation are publicly available and reproducible. Second, the analytical framework constructed in this study can be employed to monitor the real-time mental status in response to the ongoing pandemic and beyond. Third, we provide a holistic perspective of tracking and analysing mental health signals, which broadens the mental health research paradigm. We call for researchers with diverse backgrounds to extend our findings in cross-disciplinary studies, empowering global human societies to better prepare for future public health emergencies.

### CRedit authorship contribution statement

**Siqin Wang:** Conceptualization, Methodology, Software, Validation, Formal analysis, Data curation, Investigation, Writing – original draft, Visualization, Project administration, Supervision. **Xiao Huang:** Conceptualization, Methodology, Software, Validation, Formal analysis, Data curation, Writing – original draft, Visualization, Supervision. **Tao Hu:** Software, Validation, Data curation, Investigation, Visualization, Writing – review & editing. **Bing She:** Software, Validation, Data curation, Investigation, Visualization, Writing – review & editing. **Mengxi Zhang:** Investigation, Validation, Writing – original draft. **Ruomei Wang:** Software, Validation, Data curation, Investigation, Visualization. **Oliver Gruebner:** Conceptualization, Writing – review & editing. **Muhammad Imran:** Validation, Data curation, Writing – review & editing. **Jonathan Corcoran:** Conceptualization, Writing – review & editing. **Yan Liu:** Conceptualization, Writing – review & editing. **Shuming Bao:** Writing – review & editing, Supervision, Project administration.

### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Data availability

Data has provided via the links in 'Data Availability'.

### Acknowledgements

We would thank Lin Gong from University of Washington, US for helping with data collection.

### Fundin

National Science Foundation Awards #1841403.

### Data Availability

Tweet sentiment data at the individual level is accessible via CrisisNLP (<https://crisisnlp.qcri.org/tbcov>) developed by Qatar Computing Research Institute.

Tweet sentiment data at the aggregated level, and the animation and

data visualisation of tweet sentiment data at the daily level are provided in Spatial Data Lab, Center for Geographic Analysis, Harvard University ([https://projects.iq.harvard.edu/chinadatalab/global\\_twitter\\_covid19](https://projects.iq.harvard.edu/chinadatalab/global_twitter_covid19)).

Policy index data is retrieved from Oxford COVID-19 Government Response Tracker, 2020 (<https://www.bsg.ox.ac.uk/research/research-projects/covid-19-government-response-tracker>). Mental health indicators are retrieved from World Health Organization ([https://www.who.int/data/gho/data/indicators/indicator-details/GHO/hospital-beds-\(per-10-000-population\)](https://www.who.int/data/gho/data/indicators/indicator-details/GHO/hospital-beds-(per-10-000-population))), United Nations Population Division (<https://www.un.org/development/desa/pd/>) and the World Happiness Report (<https://worldhappiness.report/ed/2020/>). Source data are provided with this paper.

### Code Availability

The custom code and mathematical algorithm used for this study were generated via Python and R. They can be accessed via the project repository: ([https://projects.iq.harvard.edu/chinadatalab/global\\_twitter\\_covid19](https://projects.iq.harvard.edu/chinadatalab/global_twitter_covid19)).

### Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jag.2022.103160>.

### References

- Agarwal, A., Xie, B., Vovsha, I., Rambow, O., Passonneau, R.J., 2011. Sentiment analysis of twitter data, Proceedings of the workshop on language in social media (LSM 2011), pp. 30-38.
- Ahmed, S., 2010. The promise of happiness. Duke University Press.
- Alon-Tirosh, M., Hadar-Shoval, D., Asraf, K., Tannous-Haddad, L., Tzischinsky, O., 2021. The association between lifestyle changes and psychological distress during COVID-19 lockdown: the moderating role of COVID-related stressors. *Int. J. Environ. Res. Public Health* 18, 9695.
- Balcombe, L., De Leo, D., 2020. An integrated blueprint for digital mental health services amidst COVID-19. *JMIR Mental Health* 7, e21718.
- Belkin, M., Niyogi, P., 2001. Laplacian eigenmaps and spectral techniques for embedding and clustering. *Advances in neural information processing systems* 14.
- Blank, G., 2017. The digital divide among Twitter users and its implications for social research. *Soc. Sci. Comput. Rev.* 35, 679-697.
- Brüderl, J., Ludwig, V., 2015. Fixed-effects panel regression. *The Sage handbook of regression analysis and causal inference* 327, 357.
- Carr, M.J., Steeg, S., Webb, R.T., Kapur, N., Chew-Graham, C.A., Abel, K.M., Hope, H., Pierce, M., Ashcroft, D.M., 2021. Effects of the COVID-19 pandemic on primary care-recorded mental illness and self-harm episodes in the UK: a population-based cohort study. *Lancet Public Health* 6, e124-e135.
- Cheng, I., Heyl, J., Lad, N., Facini, G., Grout, Z., 2021. Evaluation of Twitter data for an emerging crisis: an application to the first wave of COVID-19 in the UK. *Sci. Rep.* 11, 1-13.
- Coppersmith, G., Dredze, M., Harman, C., 2014. Quantifying mental health signals in Twitter, Proceedings of the workshop on computational linguistics and clinical psychology: From linguistic signal to clinical reality, pp. 51-60.
- Cullen, W., Gulati, G., Kelly, B.D., 2020. Mental health in the COVID-19 pandemic. *QJM: An Int. J. Med.* 113, 311-312.
- Czeisler, M.E., Lane, R.L., Petrosky, E., Wiley, J.F., Christensen, A., Njai, R., Weaver, M.D., Robbins, R., Facer-Childs, E.R., Barger, L.K., 2020. Mental health, substance use, and suicidal ideation during the COVID-19 pandemic—United States, June 24–30, 2020. *Morb. Mortal. Wkly Rep.* 69, 1049.
- Díaz, F., Henríquez, P.A., 2021. Social sentiment segregation: Evidence from Twitter and Google Trends in Chile during the COVID-19 dynamic quarantine strategy. *PLoS One* 16, e0254638.
- Ewing, L.-A., Vu, H.Q., 2021. Navigating 'home schooling' during COVID-19: Australian public response on twitter. *Media Int. Australia* 178, 77-86.
- Fisher, J.R., Tran, T.D., Hammarberg, K., Sastry, J., Nguyen, H., Rowe, H., Popplestone, S., Stocker, R., Stubber, C., Kirkman, M., 2020. Mental health of people in Australia in the first month of COVID-19 restrictions: a national survey. *Med. J. Aust.* 213, 458-464.
- Hale, T., Webster, S., Petherick, A., Phillips, T., Kira, B., 2020. Oxford COVID-19 government response tracker (OxCGRT). Last updated 8, 30.
- Helliwell, J.F., Layard, R., Sachs, J., De Neve, J.E., 2020. World happiness report 2020. John F. Helliwell, R.L., Jeffrey D. Sachs, Jan-Emmanuel De Neve, Lara B. Aknin, and Shun Wang, 2021. World happiness report 2021.
- Hu, T., Wang, S., Luo, W., Yan, Y., Zhang, M., Huang, X., Liu, R., Ly, K., Kacker, V., Li, Z., 2021. Revealing public opinion towards COVID-19 vaccines using Twitter data in the United States: a spatiotemporal perspective. medRxiv.
- Hummel, S., Oetjen, N., Du, J., Posenato, E., de Almeida, R.M.R., Losada, R., Ribeiro, O., Frisardi, V., Hopper, L., Rashid, A., 2021. Mental health among medical

- professionals during the COVID-19 pandemic in eight European countries: Cross-sectional survey study. *J. Med. Internet Res.* 23, e24983.
- Hussain, A., Tahir, A., Hussain, Z., Sheikh, Z., Gogate, M., Dashtipour, K., Ali, A., Sheikh, A., 2021. Artificial intelligence-enabled analysis of public attitudes on Facebook and Twitter toward COVID-19 vaccines in the United Kingdom and the United States: Observational study. *J. Med. Internet Res.* 23, e26627.
- Imran, M., Qazi, U., Ofli, F., 2022. TBCOV: Two Billion Multilingual COVID-19 Tweets with Sentiment, Entity, Geo, and Gender Labels. *Data* 7, 8.
- Jacobson, N.C., Lekkas, D., Price, G., Heinz, M.V., Song, M., O'Malley, A.J., Barr, P.J., 2020. Flattening the mental health curve: COVID-19 stay-at-home orders are associated with alterations in mental health search behavior in the United States. *JMIR Mental Health* 7, e19347.
- Jang, H., Rempel, E., Roth, D., Carenini, G., Janjua, N.Z., 2021. Tracking COVID-19 discourse on Twitter in North America: infodemiology study using topic modeling and aspect-based sentiment analysis. *J. Med. Internet Res.* 23, e25431.
- Kouloumpis, E., Wilson, T., Moore, J., 2011. Twitter sentiment analysis: The good the bad and the omg!, Proceedings of the international AAAI conference on web and social media, pp. 538-541.
- Kwok, S.W.H., Vadde, S.K., Wang, G., 2021. Tweet topics and sentiments relating to COVID-19 vaccination among Australian Twitter users: Machine learning analysis. *J. Med. Internet Res.* 23, e26953.
- Lehoucq, R.B., Sorensen, D.C., Yang, C., 1998. ARPACK users' guide: solution of large-scale eigenvalue problems with implicitly restarted Arnoldi methods. *SIAM*.
- Levin, R., Chao, D.L., Wenger, E.A., Proctor, J.L., 2021. Insights into population behavior during the COVID-19 pandemic from cell phone mobility data and manifold learning. *Nature Comput. Sci.* 1, 588-597.
- Liu, S., Yang, L., Zhang, C., Xiang, Y.-T., Liu, Z., Hu, S., Zhang, B., 2020. Online mental health services in China during the COVID-19 outbreak. *Lancet Psychiatry* 7, e17-e18.
- Müller, M., 2007. Dynamic time warping. *Information Retrieval for Music and Motion* 69-84.
- Naseem, U., Razzak, I., Khushi, M., Eklund, P.W., Kim, J., 2021. Covidsent: A large-scale benchmark Twitter data set for COVID-19 sentiment analysis. *IEEE Trans. Comput. Social Syst.*
- Newby, J.M., O'Moore, K., Tang, S., Christensen, H., Faasse, K., 2020. Acute mental health responses during the COVID-19 pandemic in Australia. *PLoS One* 15, e0236562.
- O'Connor, R.C., Wetherall, K., Cleare, S., McClelland, H., Melson, A.J., Niedzwiedz, C.L., O'Carroll, R.E., O'Connor, D.B., Platt, S., Scowcroft, E., 2021. Mental health and well-being during the COVID-19 pandemic: longitudinal analyses of adults in the UK COVID-19 Mental Health & Wellbeing study. *Br. J. Psychiatry* 218, 326-333.
- Rahman, M.M., Ali, G.M.N., Li, X.J., Samuel, J., Paul, K.C., Chong, P.H., Yakubov, M., 2021. Socioeconomic factors analysis for COVID-19 US reopening sentiment with Twitter and census data. *Heliyon* 7, e06200.
- Ren, X., Huang, W., Pan, H., Huang, T., Wang, X., Ma, Y., 2020. Mental health during the COVID-19 outbreak in China: a meta-analysis. *Psychiatr. Q.* 1-13.
- Saleh, S.N., Lehmann, C.U., McDonald, S.A., Basit, M.A., Medford, R.J., 2021. Understanding public perception of coronavirus disease 2019 (COVID-19) social distancing on Twitter. *Infect. Control Hosp. Epidemiol.* 42, 131-138.
- Satopaa, V., Albrecht, J., Irwin, D., Raghavan, B., 2011. Finding a "kneedle" in a haystack: Detecting knee points in system behavior, 2011 31st international conference on distributed computing systems workshops. *IEEE* 166-171.
- Suratnoaji, C., Nurhadi, N., Arianto, I.D., 2020. Public opinion on lockdown (PSBB) policy in overcoming COVID-19 pandemic in Indonesia: Analysis based on big data twitter. *Asian J. Publ. Opin. Res.* 8, 393-406.
- Talevi, D., Soccì, V., Carai, M., Carnaghi, G., Faleri, S., Trebbi, E., di Bernardo, A., Capelli, F., Pacitti, F., 2020. Mental health outcomes of the COVID-19 pandemic. *Riv. Psichiatr.* 55, 137-144.
- Tan, E.J., Meyer, D., Neill, E., Phillipou, A., Toh, W.L., Van Rheenen, T.E., Rossell, S.L., 2020. Considerations for assessing the impact of the COVID-19 pandemic on mental health in Australia. *Aust. N. Z. J. Psychiatry* 54, 1067-1071.
- Ting, R.-S.-K., Yong, Y.-Y.-A., Tan, M.-M., Yap, C.-K., 2021. Cultural responses to COVID-19 pandemic: religions, illness perception, and perceived stress. *Front. Psychol.* 12.
- Torres-Reyna, O., 2007. Panel data analysis fixed and random effects using Stata. *Data & Statistical Services*, Princeton University, p. 112 v. 4.2).
- Tran, T.D., Hammarberg, K., Kirkman, M., Nguyen, H.T.M., Fisher, J., 2020. Alcohol use and mental health status during the first months of COVID-19 pandemic in Australia. *J. Affect. Disord.* 277, 810-813.
- Tufekci, Z., 2014. Big questions for social media big data: Representativeness, validity and other methodological pitfalls, Eighth international AAAI conference on weblogs and social media.
- Twitter, 2020. Twitter API.
- United Nations Population Division, 2020.
- Van Rheenen, T.E., Meyer, D., Neill, E., Phillipou, A., Tan, E.J., Toh, W.L., Rossell, S.L., 2020. Mental health status of individuals with a mood-disorder during the COVID-19 pandemic in Australia: initial results from the COLLATE project. *J. Affect. Disord.* 275, 69-77.
- Venna, J., Kaski, S., 2001. Neighborhood preservation in nonlinear projection methods: An experimental study. In: *International Conference on Artificial Neural Networks*. Springer, pp. 485-491.
- Wang, S., Huang, X., Hu, T., Zhang, M., Li, Z., Ning, H., Corcoran, J., Khan, A., Liu, Y., Zhang, J., 2022. The times, they are a-changin': tracking shifts in mental health signals from early phase to later phase of the COVID-19 pandemic in Australia. *BMJ Glob. Health* 7, e007081.
- Wang, Y., Shi, L., Que, J., Lu, Q., Liu, L., Lu, Z., Xu, Y., Liu, J., Sun, Y., Meng, S., 2021. The impact of quarantine on mental health status among general population in China during the COVID-19 pandemic. *Mol. Psychiatry* 1-10.
- Wang, J., Zhou, Y., Zhang, W., Evans, R., Zhu, C., 2020. Concerns expressed by Chinese social media users during the COVID-19 pandemic: Content analysis of sina weibo microblogging data. *J. Med. Internet Res.* 22, e22152.
- World Health Organization, 2020a. The global health observatory. Indicators. Hospital beds (per 10,000 population).
- World Health Organization, 2020b. Timeline: WHO's COVID-19 response 2020. World Health Organization.
- World Health Organization, 2020c. WHO issues its first emergency use validation for a COVID-19 vaccine and emphasizes need for equitable global access.
- Zhou, J., Zogan, H., Yang, S., Jameel, S., Xu, G., Chen, F., 2021. Detecting community depression dynamics due to COVID-19 pandemic in Australia. *IEEE Trans. Comput. Social Syst.* 8, 982-991.