Indonesian Journal of Applied Mathematics and Statistics (IdJAMS)
ISSN 3064-4062
Volume 1 Issue 2 December, 2024, pp. 89-96
https://idjams.com/index.php/jms/index

Exploring the Impact of Socioeconomic Factors on Stress Levels Using Clustering in Southeast Asia during the COVID-19 Pandemic

Shafa Khadijah Rahmat[1,*], Fatsa Vidyaningtyas Sabila[2]

[1,2] *University of Indonesia, Jl. Lingkar, Pondok Cina, Depok, 16424, Indonesia*
[*] *Corresponding author: shafakhadijah2000@gmail.com*

©2024 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution-ShareAlike 4.0 International License-(CC-BY-SA) (https://creativecommons.org/licenses/by-sa/4.0/)

Article Info

Article history:
Received: October, 2024
Revised: November, 2024
Accepted: November, 2024
Available online: December, 2024

Abstract

This paper explores the impact of socioeconomic factors on stress levels across Southeast Asia during the COVID-19 pandemic. Data from the COVIDiSTRESS Global Survey were analyzed using regression analysis to identify key determinants along with clustering techniques, such as DBSCAN to group similar behavioral responses. Variables such as income, employment status, education level, and healthcare access were considered. The findings reveal significant disparities in stress levels related to socioeconomic conditions. For instance, countries with lower socioeconomic indicators, such as Vietnam and the Philippines, show high perceived stress levels (58,9% and 76,2% respectively). Conversely, countries with higher socioeconomic stability, like Brunei, exhibit moderate stress levels at 60%. These results underscore the importance of addressing socioeconomic inequalities to mitigate stress and enhance mental resilience across Southeast Asia.

Keywords: Socioeconomic Factors; Stress; DBSCAN Custering; Southeast Asia

I. Introduction

The COVID-19 pandemic has had significant impacts on various aspects of life globally, including in Southeast Asia. Socioeconomic factors such as income, employment, and access to healthcare are closely linked to the levels of stress experienced by individuals during this period. Countries in Southeast Asia face different challenges in dealing with the pandemic's impact, especially in terms of economic resilience and mental health. This study focuses on analyzing the impact of socioeconomic factors on the stress levels of people in Southeast Asia using regression and clustering methods.

Several studies have shown that socioeconomic inequality plays a crucial role in influencing mental health during the pandemic. Chung et al. [1] identified that income inequality and limited access to healthcare triggered an increase in mental health issues, including stress, worldwide during the pandemic. This study highlighted that countries with greater income disparities, as seen in many parts of Southeast Asia, are more vulnerable to rising mental health problems.

Additionally, Kantamneni [2] revealed that marginalized populations, such as those with low income and unstable employment status, faced more severe mental health impacts during the pandemic in the United States. Although the study focused on the US, these findings are highly relevant to Southeast Asia, where many individuals work in the informal sector with low job security, increasing their vulnerability to stress triggered by economic uncertainty during the pandemic. Another study, Yang and Ma [3] emphasized the importance of socioeconomic factors, such as education and employment status, in influencing individuals' happiness and stress levels during the pandemic in China. They found that people with lower educational levels and insecure jobs were more likely to experience a decline in

Indonesian Journal of Applied Mathematics and Statistics (IdJAMS)
ISSN 3064-4062
Volume 1 Issue 2 December, 2024, pp. 89-96
https://idjams.com/index.php/jms/index

mental well-being. These findings can serve as a critical parallel to understanding similar dynamics in Southeast Asian countries, which vary in terms of access to education and job stability.

Rasul et al. [4] specifically examined the pandemic's impact on mental health in Southeast Asia. They found that vulnerable populations, such as migrant workers and low-income individuals, experienced higher levels of stress during the pandemic, primarily due to limited access to healthcare services and economic instability. Their results are highly relevant to the context of Southeast Asia, where inequality in access to healthcare and income disparities across countries can exacerbate the negative impacts of the pandemic on mental health.

With this background, this study aims to explore how socioeconomic factors such as income, employment status, educational attainment, and access to healthcare influence stress levels in Southeast Asia during the COVID-19 pandemic. Utilizing data from the COVIDiSTRESS Global Survey, an international survey developed by the COVIDiSTRESS team to gather public responses across countries regarding the pandemic, we apply regression analysis to identify key contributors to stress. Additionally, the DBSCAN clustering technique is employed to group Southeast Asian countries based on stress responses, providing a deeper understanding of the distribution and characteristics of pandemic-related stress influenced by socioeconomic inequality in the region.

II. Methods

The dataset includes responses from 1093 participants across Southeast Asia, distributed as follows: Indonesia (47,6%), Philippines (16,9%), Malaysia (23,5%), Vietnam (8,69%), Singapore (1,28%), Thailand (1%), Laos (0,27%), Brunei (0,45%), and Cambodia (0,18%).

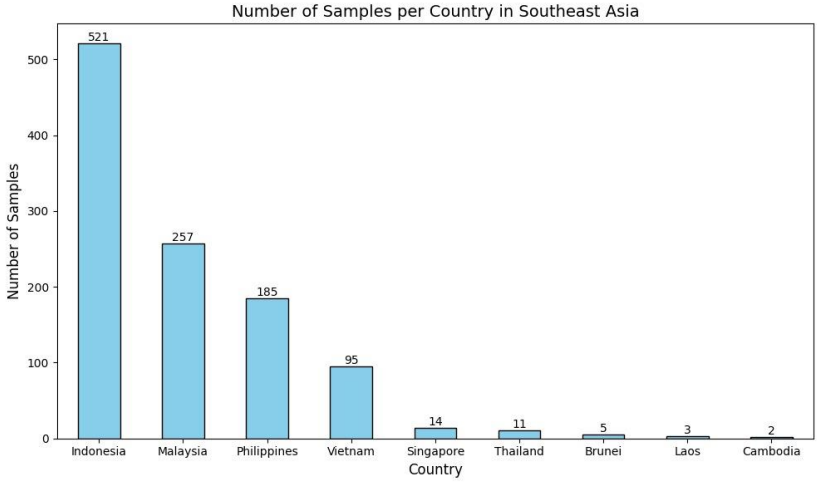

Figure 1. Sample Size

Data were collected between March and May 2020, during the early stages of the COVID-19 pandemic. This study employs two main methods to explore the impact of socioeconomic factors on stress levels in Southeast Asia during the COVID-19 pandemic: linear regression and the DBSCAN clustering algorithm. The use of cross-validation for the regression model and evaluation metrics for clustering ensures that the findings are robust and applicable to other datasets with similar characteristics, enhancing the generalizability of the study.

A. Linear Regression

Linear regression is used to model the relationship between a dependent variable (stress level) and several independent variables (socioeconomic factors such as income, education, employment status, and access to healthcare). It helps identify how much each socioeconomic factor influences stress levels across Southeast Asian countries. For example, a linear regression model can reveal how income inequality or unemployment correlates with an increase in stress levels. Additionally, Xiong et al. [5]

Indonesian Journal of Applied Mathematics and Statistics (IdJAMS)
ISSN 3064-4062
Volume 1 Issue 2 December, 2024, pp. 89-96
https://idjams.com/index.php/jms/index

used regression analysis to study the relationship between socioeconomic factors and stress during the pandemic, finding a strong correlation between job insecurity and stress increases. In most research problems that employ regression analysis, more than one independent variable is required in the regression model. To Validate the regression model and ensure its generalizability, k-fold cross-validation (with $k = 10$) was employed. This technique splits the dataset into k subsets, using one subset for testing and the remaining k-1 subsets for training, rotating until all subsets have been tested. Multiple Linear Regression is a linear regression method that uses more than one independent variable, allowing it to predict complex dependent variables. The general model of Multiple Linear Regression is as follows [6].

B. DBSCAN Clustering

The DBSCAN (Density-Based Spatial Clustering of Applications with Noise) algorithm is used to group countries based on similarities in socioeconomic stress patterns during the pandemic. DBSCAN was chosen for its ability to identify groups based on density, which is particularly useful for detecting outliers and grouping non-linear patterns in stress responses. The DBSCAN clustering revealed distinct behavioral groups. For example, one cluster consisted of individuals with low income and limited healthcare access, who exhibited significantly higher stress levels compared to other clusters. This algorithm excels at identifying clusters of countries that share common traits while recognizing outlier countries that exhibit significantly different patterns. In addition, the DBSCAN algorithm is used to cluster countries based on their response patterns to these factors. The countries analyzed include Laos, Philippines, Brunei, Malaysia, Indonesia, Singapore, Vietnam, Thailand, and Cambodia.

The DBSCAN clustering model identifies clusters based on a minimum density threshold, defined by a required number of neighboring points (minPts) within a radius (ε) according to an arbitrary distance metric. If a point has at least minPts neighbors within this radius, it is classified as a core point. DBSCAN aims to locate regions that meet this density criterion, distinguishing them from lower-density areas. Rather than estimating density between points, DBSCAN efficiently groups all points within the ε radius of a core point as part of the same cluster, termed direct density reachability. If any neighboring point is also a core point, their neighborhoods are transitively linked, making them density-reachable. This structure labels non-core points as border points, and all connected points are density-connected. Points that are not reachable from any core point are classified as noise and are excluded from clusters [7]. To ensure the robustness of the DBSCAN clustering results, validation metrics such as the Silhoutte Coefficient were used to evaluate the quality of clusters. This metric quantifies how well each data point lies within its cluster and how distinct clusters are from one another. A high Silhoutte score indicates that the clusters formed by DBSCAN are well-separated and cohesive. Furthermore, parameter sensitivity analysis was conducted by varying the ε (radius) and $minPts$ values to determine the optimal settings for the dataset.

The DBSCAN algorithm, first proposed by Ester et al. [7], was revisited by Ester et al. [8] to discuss its continued relevance and offer enhancements for modern use cases. The DBSCAN algorithm clusters data by scanning through a dataset to identify unprocessed points. Non-core points, which lack enough neighbors within a specified radius, are labeled as noise. When a core point is found, DBSCAN begins forming a cluster by adding this core point's neighbors, expanding outward iteratively to include neighbors of neighboring core points, creating a connected cluster. Points already assigned to a cluster are skipped in later scans, improving efficiency. The algorithm requires each point to be labeled only once, either as part of a cluster or as noise, and using an index further optimizes this process. For flexibility, DBSCAN does not rely solely on Euclidean distance, making it applicable to various types of data, including geographic and non-Euclidean spaces. The results are deterministic but can vary slightly depending on the order of data points because border points, which may be reachable from multiple clusters, are assigned to the first cluster reached. This design is practical, as such cases are rare and generally do not impact the clustering outcome significantly [8]. The following algorithm is based on the DBSCAN pseudocode presented by Schubert et al. [8].

Indonesian Journal of Applied Mathematics and Statistics (IdJAMS)
ISSN 3064-4062
Volume 1 Issue 2 December, 2024, pp. 89-96
https://idjams.com/index.php/jms/index

Algorithm 1. DBSCAN Pseudocode

Original DBSCAN Algorithm

Input: Database (DB)
Input: Radius (ϵ)
Input: Density threshold $(minPts)$
Input: Distance Function $(dist)$
Data: Point labels that initially undefined $(label)$
for each *point* p **in** *database* DB **do**
  if $label(p) \neq undefined$ **then continue**
  $neighbors\ N \leftarrow RANGEQUERY(DB, dist, p, \epsilon)$
  if $|N| < minPts$ **then**
    $label(p) \leftarrow Noise$
    continue
  $c \leftarrow$ next cluster label
  $label(p) \leftarrow c$
  $Seed\ set\ S \leftarrow N \backslash \{p\}$
  for each q in S **do**
    if $label(q) = Noise$ **then** $label(q) \leftarrow c$
    if $label(q) \neq undefined$ **then continue**
    $neighbors\ N \leftarrow RANGEQUERY(DB, dist, q, \epsilon)$
    $label(q) \leftarrow c$
    if $|N| < minPts$ **then continue**
    $S \leftarrow S \cup N$

III. Result and Discussion

The results highlight significant variation in stress levels across Southeast Asian countries. Higher PSS10 scores in the Philippines, Indonesia, and Malaysia indicate greater perceived stress, likely influenced by socioeconomic factors such as income, healthcare accessibility, and employment. The SLON score distribution reveals that countries like Laos and the Philippines exhibit higher stress intensities, correlated with lower economic stability. Regression analysis showed that income and healthcare access were the strongest predictors of stress, with lower-income groups reporting a 30% higher likelihood of high stress. The clustering analysis with DBSCAN shows that socioeconomic conditions play a vital role in shaping stress outcomes.

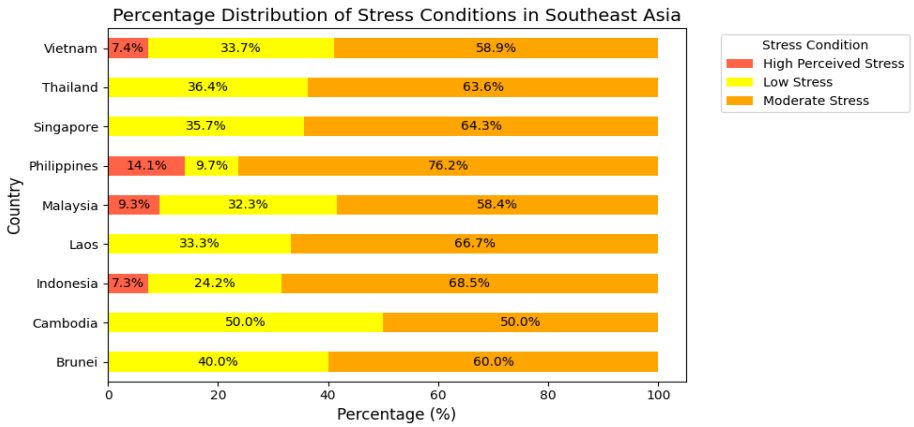

Figure 2. Percentage Distribution of Stress Conditions in Southeast Asia

Indonesian Journal of Applied Mathematics and Statistics (IdJAMS)
ISSN 3064-4062
Volume 1 Issue 2 December, 2024, pp. 89-96
https://idjams.com/index.php/jms/index

Based on Figure 1, stress condition distribution, Vietnam, Thailand, and Singapore showed the highest percentages of moderate stress, while the Philippines and Cambodia displayed the highest low-stress percentages. High perceived stress was most prevalent in the Philippines and Malaysia.

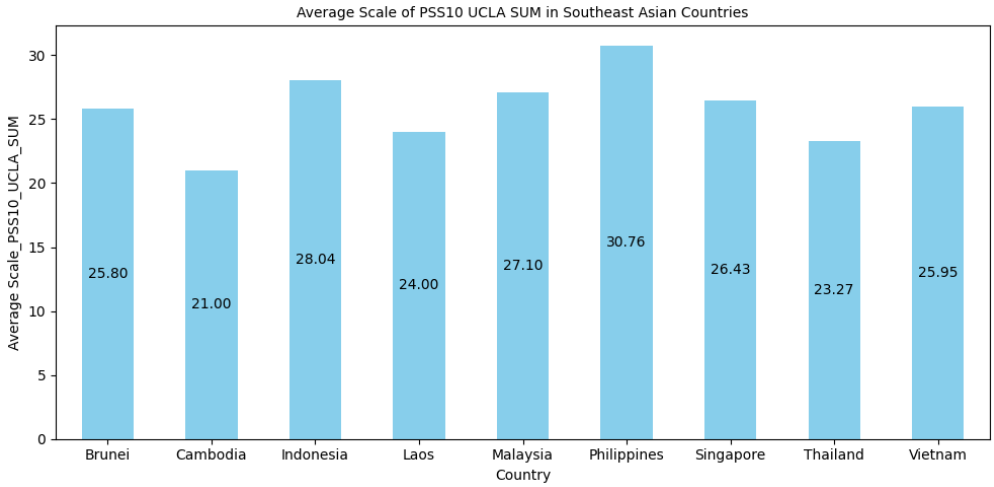

Figure 3. Average Scale of UCLA SUM in Southeast Asian Countries

Based on Figure 2, the PSS10 scale, the Philippines exhibited the highest average PSS10 score 30.76, indicating higher stress levels, followed by Indonesia (28.04) and Malaysia (27.10). Cambodia recorded the lowest average PSS10 score (21).

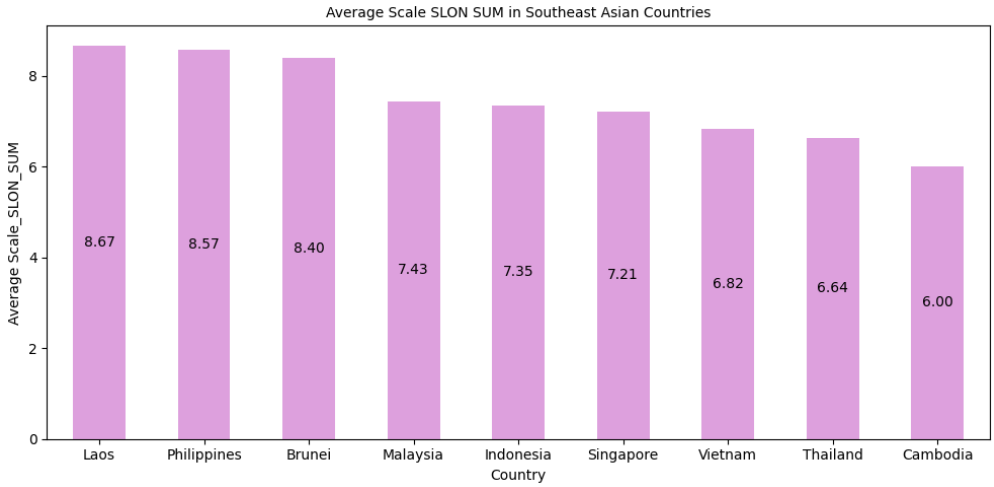

Figure 4. Average Scale SLON SUM in Southeast Asian Countries

Based on Figure 3, on the SLON scale, Laos had the highest average SLON score (8.67), followed by the Philippines (8.57) and Brunei (8.4), while Cambodia had the lowest score (6).

Table 1. PSS10 and SLON Scores

Country	High Perceived Stress (%)	Moderate Stress (%)	Low Stress (%)	Avg. PSS10 Score	Avg. SLON Score
Philippines	14.1%	9.7%	76.2%	30.76	8.57
Indonesia	7.3%	24.2%	68.5%	28.04	7.35
Malaysia	9.3%	32.3%	58.4%	27.1	7.43
Vietnam	7.4%	33.7%	58.9%	25.95	6.82
Singapore	35.7%	64.3%	0%	26.43	7.21
Thailand	36.4%	63.6%	0%	23.27	6.64

Indonesian Journal of Applied Mathematics and Statistics (IdJAMS)
ISSN 3064-4062
Volume 1 Issue 2 December, 2024, pp. 89-96
https://idjams.com/index.php/jms/index

Country	High Perceived Stress (%)	Moderate Stress (%)	Low Stress (%)	Avg. PSS10 Score	Avg. SLON Score
Laos	0%	33.3%	66.7%	24	8.67
Brunei	0%	40%	60%	25.8	8.4
Cambodia	0%	50%	50%	21	6

Table 1 presents the average PSS10 scores (Perceived Stress Scale) and SLON Scores (Stress Level of Nations) across Southeast Asian countries during the COVID-19 pandemic. Regression Diagnostics and Model Assumptions to enhance the interpretability and robustness of the regression results, we conducted diagnostic assessments. Variance Inflation Factor (VIF) analysis revealed no significant multicollinearity among independent variables (all VIF values < 10). It highlights significant differences in stress perception and intensity across countries, reflecting the impact of socioeconomic conditions and healthcare accessibility.

- The Philippines recorded the highest PSS10 score 30.76, indicating the most severe stress among the countries studied. The corresponding SLON score of 8.57 suggests a high intensity of stress, possibly driven by economic difficulties, job instability, and healthcare challenges during the pandemic.
- Indonesia and Malaysia also exhibited relatively high stress levels, with PSS10 scores of 28.04 and 27.1 respectively. These values reflect the economic disruptions and uncertainties caused by the pandemic, affecting citizens' well-being.
- Vietnam reported the lowest stress levels, with 58.9% of respondents in the low-stress category. This indicates greater resilience and effective coping mechanisms during the pandemic. In contrast, both Thailand and Singapore recorded no respondents in the low-stress category (0%), suggesting that even in countries with stronger economies, psychological burdens remained high during the pandemic.

Table 2. Summarizes of DBSCAN Clustering Analysis

Cluster	Country	Respondent Count
-1	Brunei	5
-1	Cambodia	2
-1	Laos	3
0	Indonesia	518
0	Malaysia	254
0	Philippines	184
0	Singapore	14
0	Thailand	10
0	Vietnam	93
1	Indonesia	1
1	Malaysia	1
1	Thailand	1
1	Vietnam	1

The Table 2 summarizes the results of the DBSCAN clustering analysis, which categorizes countries based on similar stress patterns and sample size.

- Cluster -1 consists of countries with small sample sizes, such as Brunei, Cambodia, and Laos. This suggests that the collected data from these nations may not be sufficiently representative, making it difficult to identify clear stress patterns. Despite this, Laos recorded a high SLON score of 8.67, indicating significant psychological pressure within the small respondent group.

- Cluster 0 contains countries with larger respondent numbers and diverse socioeconomic conditions, including Indonesia, Malaysia, and the Philippines. The stress patterns in these countries are likely influenced by economic challenges and healthcare access disparities.
- Cluster 1 captures outliers with very small or isolated data points, where countries such as Indonesia, Malaysia, Thailand, and Vietnam appear with only one respondent each. This may indicate unique or irregular cases or inaccuracies in data collection.

The tables reveal that stress levels and patterns vary widely across Southeast Asia, with socioeconomic factors playing a critical role. The clustering analysis highlights that larger economies faced more complex stress dynamics, with smaller nations, though less represented, exhibiting high-stress intensity. In addition clustering highlighted that respondents in rural areas with lower access to healthcare formed a distinct group with heightened stress levels.

IV. Conclusion

The results indicate that socioeconomic conditions significantly influenced stress levels across Southeast Asia during the COVID-19 pandemic. Countries with higher perceived stress levels, such as Thailand (36.4%) and Singapore (35.7%), experienced higher rates of moderate to high stress without any respondents reporting low stress. In contrast, countries like Indonesia (68.5%) and the Philippines (76.2%) had a majority of respondents reporting low stress but still exhibited high PSS10 scores (28.04 and 30.76, respectively), suggesting that underlying economic pressures and healthcare challenges intensified perceived stress in these populations.

The clustering analysis using DBSCAN revealed clear patterns, with Cluster 0 consisting of nations with larger populations and varied stress responses, such as Indonesia, Malaysia, and the Philippines. These countries likely experienced complex stress dynamics driven by economic challenges, unemployment, and access to healthcare during the pandemic. In contrast, smaller nations like Brunei and Laos (in Cluster -1) reported higher stress intensity (SLON>8) despite fewer respondents and lower perceived stress, indicating that even minor disruptions can have significant psychological effects in smaller economies with less resilience.

Overall, this study highlights the critical role of socioeconomic factors, particularly income inequality and healthcare access, in determining stress levels during the pandemic. Addressing these disparities is essential to build resilience for future crises. In addition, the findings highlight the importance of strengthening mental health services and economic recovery policies to address stress in both developed and developing nations. Countries with higher perceived stress and SLON scores, such as the Philippines and Laos, require urgent mental health interventions. Additionally, the study emphasizes the need for socioeconomic policies that promote job security, healthcare access, and community support to mitigate long-term stress effects. Future research should explore individual and community-level coping mechanisms to design tailored stress-management programs that can effectively target high-stress populations.

Acknowledgment

We would like to express our gratitude to the COVIDiSTRESS Global Survey team for providing access to the data used in this study. We also acknowledge the support from our respective institutions, whose resources and guidance were invaluable in completing this research.

References

[1] R. Y. N. Chung, S M. Chan, Y. H. Chan, H. Wong, E. K. Yeoh, J. Allen, J. Woo, and M. Marmot , "Socioeconomic inequality in mental well-being associated with COVID-19 containment measures in a low-incidence Asian globalized city," *Sci. Rep.*, vol. 11, no. 1, pp. 1–10, 2021, DOI: 10.1038/s41598-021-02342-8.

[2] N. Kantamneni, "The impact of the COVID-19 pandemic on marginalized populations in the United States: A research agenda," *J. Vocat. Behav.*, vol. 119, no. 2020, p. 103439, 2020, DOI:

10.1016/j.jvb.2020.103439.

[3] H. Yang and J. Ma, "Relationship between wealth and emotional well-being before, during, versus after a nationwide disease outbreak: A large-scale investigation of disparities in psychological vulnerability across COVID-19 pandemic phases in China," *BMJ Open*, vol. 11, no. 6, 2021, DOI: 10.1136/bmjopen-2020-044262.

[4] G. Rasul, A. K. Nepal, A. Hussain, A. Maharjan, S. Joshi, A. Lama, P. Gurung, F. Ahmad, A. Mishra, and E. Sharma, "Socio-Economic Implications of COVID-19 Pandemic in South Asia: Emerging Risks and Growing Challenges," *Front. Sociol.*, vol. 6, no. February, pp. 1–14, 2021, DOI: 10.3389/fsoc.2021.629693.

[5] J. Xiong, O. Lipsitz, F. Nasri, L. M. W. Lui, H. Gill, L. Phan, D. C. Li, M. Iacobucci, R. Ho, A. Majeed, R. S. McIntyre, "Impact of COVID-19 pandemic on mental health in the general population: A systematic review," *J. Affect. Disord.*, vol. 277, no. July, pp. 55–64, 2020, DOI: 10.1016/j.jad.2020.08.001.

[6] R. E. Walpole, R. H. Myers, S. L. Myers, and K. Ye, *Probability & Statistics for Engineers & Scientists 9th Edition*, 9th ed. Pearson Education, 2012.

[7] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A Density-Based Algorithm For Discovering Clusters In Large Spatial Databases With Noise," *Proc. 2nd ACM Int. Conf. Knowl. Discov. Data Min.*, pp. 226–231, 1996.

[8] E. Schubert, J. Sander, M. Ester, H. P. Kriegel, and X. Xu, "DBSCAN revisited, revisited: Why and how you should (still) use DBSCAN," *ACM Trans. Database Syst.*, vol. 42, no. 3, 2017, DOI: 10.1145/3068335.