VOL. 22 (2) (2025), 35-44, p-ISSN: 693-7554, e-ISSN:2654-3990

The Similarity of COVID-19-Related Profiles and Pandemic Conditions in ASEAN Countries

Elis Khatizah

Mathematics Study Program, School of Data Science, Mathematics and Informatics, IPB University, West Java, 16680, Indonesia

Abstract

This paper examines the potential relationship between countries' profiles and COVID-19 conditions in 10 ASEAN countries, focusing on new confirmed cases and deaths. We initially clustered the countries based on demographic profiles and then observed the pandemic situation within each cluster. The samples of genome sequence, particularly the Omicron variant collected from each country, are also clustered to investigate the spread of the virus. Furthermore, we created a simple algorithm to identify countries with profiles similar to ASEAN nations. Our findings show that countries with similar demographic profiles and geographically proximate are more likely to experience similar pandemic conditions even though they have no identical genome sequence of the virus. Five of the 10 ASEAN countries have similar demographic profiles to countries within the ASEAN region and as assumed, experienced similar COVID-19 situations. Conversely, the remaining ASEAN countries, which share profiles with countries outside the region, demonstrate fairly different COVID-19 conditions, particularly with regard to the timing of the spread of the virus. With comparable resources at hand, insights from countries with similar profiles provide valuable information for effective pandemic management efforts in the future.

Keywords: *ASEAN profiles; clustering; COVID-19; demographic; pandemic management*

1. Introduction

The persistence of the COVID-19 pandemic worldwide for over three years has significantly impacted all aspects of human life. From the onset and throughout the pandemic, extensive research has been conducted to address various factors, including accurate symptom detection and practical measures such as lockdowns and social distancing to prevent transmission [1], [2]. Subsequently, vaccination programs have become a universal policy in countries across the globe [3], accompanied by the concept of the 'new normal,' where humans have learned to coexist with the virus as its initial severity diminishes. The pandemic's effects, particularly in economic and social matters, have also drawn the attention of scholars who analyze the impact and explore strategies to support human life sustainability in the post-COVID era [4], [5].

Observing the COVID-19 situation and understanding the variations in response among different countries and their outcomes is crucial [6], in line with the motivation to learn from past experiences for a better future. A study conducted in Italy highlights that COVID-19-related deaths in the country may have been higher compared to other similarly affected nations due to its unique demographic and socioeconomic profile [7]. This unique profile influences the focus of COVID-19 research and observations conducted merely in certain regions, which may limit the generalizability of findings to others. For example, the World Health Organization (WHO) categorized global regions when reporting confirmed cases and fatalities related to COVID-19 on June 9, 2020 [8].

Two critical factors identified for successful pandemic management are state-centric policies and socio-demographic factors, such as age [9]. In addition to areas with unique profiles experiencing unique situations, areas sharing similar profiles may potentially undergo similar situations. Even if conditions differ, where some sites fare

*Corresponding author. *E-mail address:* elis_khatizah@apps.ipb.ac.id
Received: 5 May 2025, Accepted: 31 July 2025 and available online 31 July 2025
DOI: https://doi.org/10.33751/komputasi.v19i2.5260

better or worse, they can still provide valuable insights to each other on handling the pandemic by leveraging their similar resource profiles. Motivated by this assumption, this paper aims to investigate the relationship between countries' profiles and the number of confirmed cases and deaths over three years as a valuable metric to observe the COVID-19 situation in 10 ASEAN countries and their comparable nations. To the best of our knowledge, no existing literature has examined this significant relationship. This study is expected to provide meaningful insights for the 10 ASEAN countries to overcome future pandemics based on their potential resources and track record in dealing with the three- year COVID-19 pandemic.

2. Methods

2.1 Data

Our dataset comprised COVID-19 information and countries' profiles, downloaded from the Our World in Data (OWID) repository [10]. Utilizing this single open-source data repository ensured better transparency and consistency of data management, analysis, and interpretation [11]. To satisfy the requirement of complete information on countries' profiles, the dataset was extracted into 159 countries worldwide, including the ten focused ASEAN countries: Brunei, Cambodia, Indonesia, Laos, Malaysia, Myanmar, Philippines, Singapore, Thailand, and Vietnam. For examining the COVID-19 situation, the extracted data includes daily new confirmed cases and daily new confirmed deaths per million population, covering three years from March 1, 2020, to March 1, 2023. The country's profile is represented by 11 features: population, population density, diabetes prevalence, cardiovascular death rate, median age, age 65 or older, age 70 or older, life expectancy, human development index, GDP per capita, and hospital beds per thousand people. To gather additional information about the viruses spreading within ASEAN countries, we also downloaded the genome sequences of the viruses from the GISAID website [12]. It is important to note that SARS-CoV-2, the virus caused COVID-19, is an RNA virus. However, we use its DNA sequence data, which is obtained by reverse transcribing the virus' RNA sequence and storing it in the database. We select 10 DNA sequences of the Omicron variant per country which were collected in October 2022.

2.2 Experiment set up

We conduct two stages to observe the COVID-19 situation in 10 ASEAN countries and their comparable nations. First, we utilize a dataset of country profiles to examine the similarity of profiles among the 10 ASEAN countries. We also employ these COVID-19-related profile similarities to identify other countries worldwide that exhibit similarities to each ASEAN member. In addition to these processes, we used Biopython's built-in functions to cluster the DNA sequences, highlighting the genetic similarities of viruses that were spreading within the countries. In the second step, we visualize the number of COVID- 19 new confirmed cases and deaths based on the countries' profile similarity. Figure 1 depicts the experimental diagram.

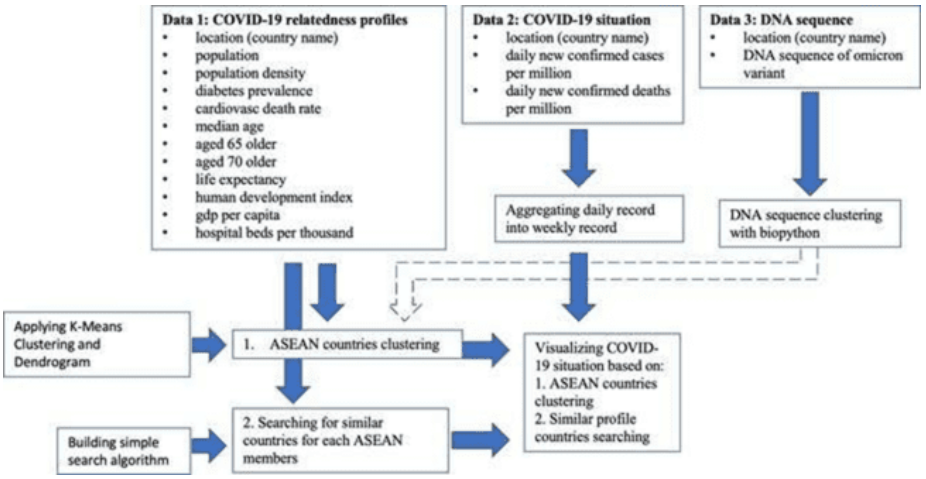


Figure 1. Experimental diagram for observing COVID-19 situation in 10 ASEAN countries and their comparable nations

To detail, in the first stage, dendrogram and K-Means clustering techniques are employed to group the 10 ASEAN countries based on their profiles. Simultaneously, we created a simple algorithm to determine a similarity profile value between a chosen country of interest and other countries. Before obtaining this value, we rescale the numerical quantity of countries' profiles to a standard range and take the square of the difference between each standardized profile of the two countries. The similarity profile value was easily calculated from the sum of these squares. This value indicates the level of similarity, with smaller values indicating a higher degree of similarity. By sorting this value, we considered the three most similar countries and selected one to visualize their COVID-19 situation compared to each ASEAN member country. Figure 2 illustrates the pseudocode for this simple algorithm. We considered the top three similar countries and selected one of them to visualize their COVID-19 situation in comparison with each ASEAN member country.

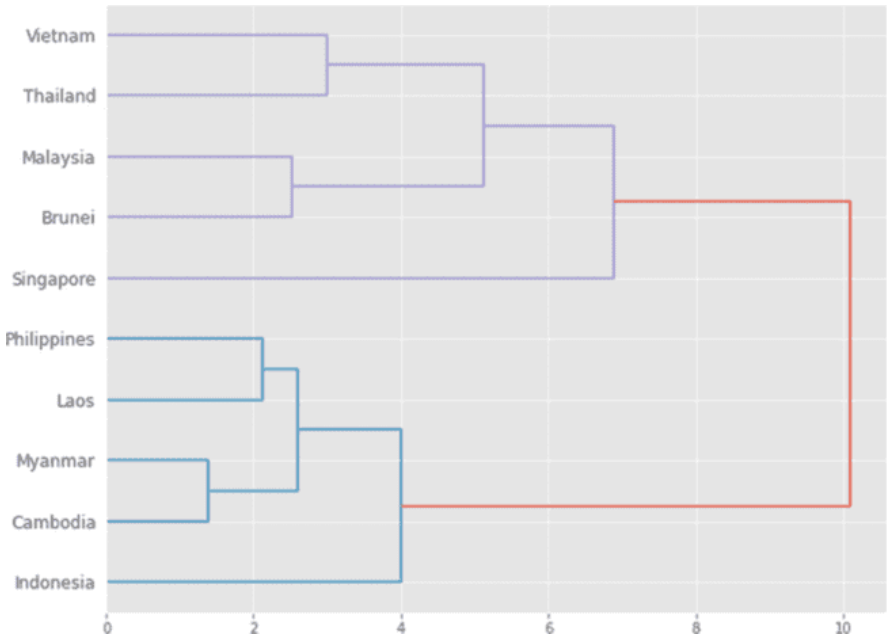
```
BEGIN
    Read data file
    similar_countries=[]
    similarity_matrix=[]
    similarity_index=[]
    INPUT interested_country
    FOR each row in data
        IF interested_country in country_name of the row
            home = interested_country_profile array
        END IF
    END FOR
    FOR each row in data
        away = country_name array
        s = sum(home-away)**2
        IF interested_country not in row
          Add  country_name to similar_countries
          Add away to similarity_matrix
          Add s to similarity_index
        END IF
     END FOR
     Dict_similar_countries1= dictionary of similar_countries and similarity index
     Dict_similar_countries2= dictionary of similar_countries and similarity matrix
     SortDict_similar_countries1= sort of Dict_similar_countries1 based similarity_index
    print the first three SortDict_similar_countries1 key
END
```

Figure 2. The pseudocode of simple algorithm for finding the most similar profile country

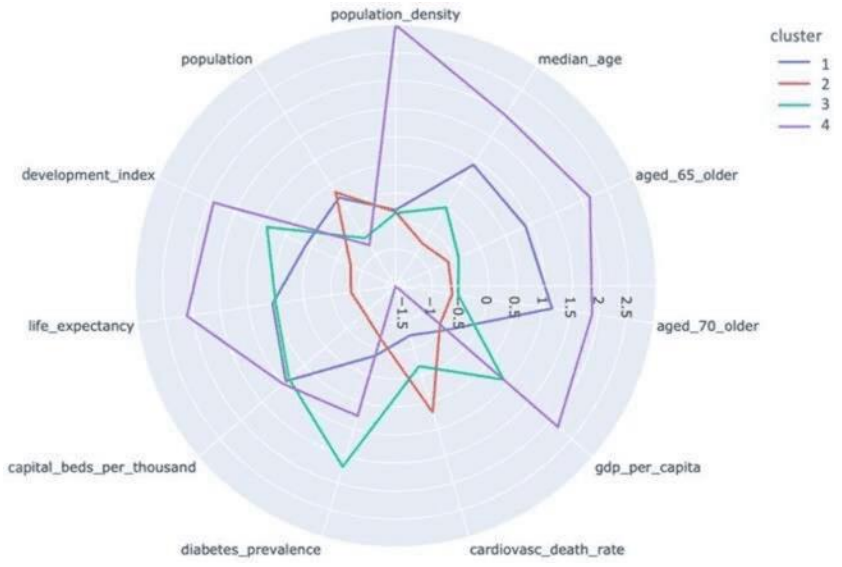
3. Results and Discussion

3.1 COVID-19 situation among 10 ASEAN countries

By using a dendrogram, in Figure 3(a) we can conveniently select the desired number of clusters for grouping the 10 ASEAN countries based on their profiles. However, the K-means clustering technique suggests that 4 clusters appear to be the optimal choice where the character for each cluster is depicted in the radar chart of Figure 3(b). With the selection of 4 clusters, Thailand and Vietnam are placed in Cluster 1, while Cluster 2 comprises Cambodia, Indonesia, Laos, Myanmar, and the Philippines. Brunei and Malaysia form Cluster 3, and Singapore stands alone in Cluster 4. As the cluster with the most members, Cluster 2 in Figure 3(b) exhibits distinctive features, including a high population count and cardiovascular death rate, as well as lower values for the human development index, life expectancy, hospital bed availability, and the number of elderly individuals. In contrast, Cluster 4 demonstrates the lowest population count and cardiovascular death rate among the clusters, accompanied by the highest values for most variables (excluding diabetes prevalence).



(a)



(b)

Figure 3. The result of clustering 10 ASEAN countries in the form of (a) a dendrogram and (b) a radar chart

To observe the COVID-19 situation, we visualize the daily new confirmed cases and deaths per million population from the dataset. Since these numbers fluctuate on a daily basis over a span of 3 years, we opt for a more convenient graphic representation by plotting the weekly values instead of daily values. Figure 4 illustrates the weekly new confirmed cases and deaths for each cluster.

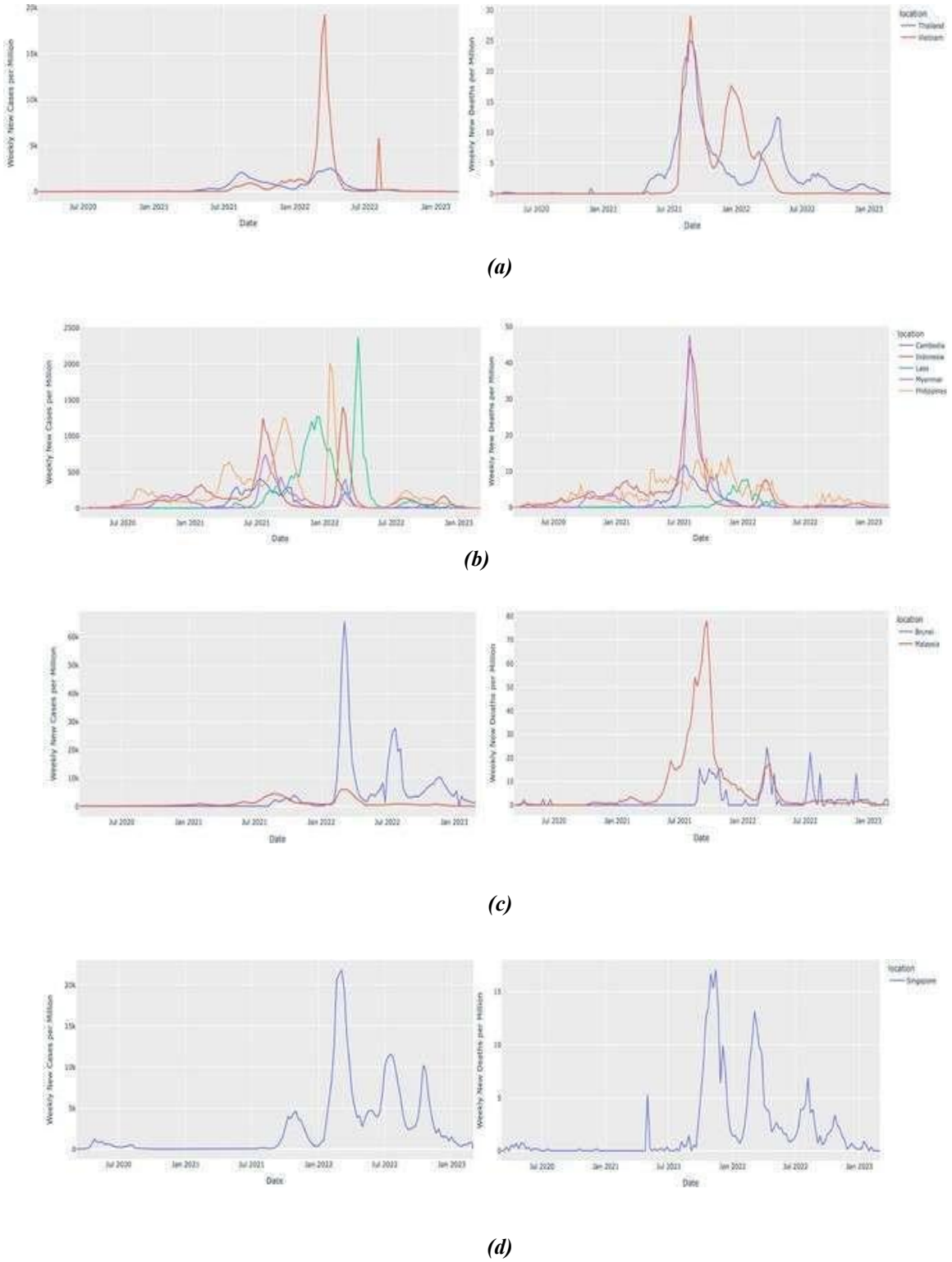


Figure 4. *The weekly new confirmed cases and deaths for (a) Cluster 1, (b) Cluster 2, (c) Cluster 3, and (d) Cluster 4*

Based on Figure 4, we observe variations in the maximum number of weekly confirmed cases and deaths among the clusters. However, all countries have a similar pattern regarding the peak period for both numbers, where the peak period for weekly new confirmed cases mostly occurred in February - March 2022, and weekly

new confirmed deaths surged in August - October 2021. During this period, Brunei experienced the highest number of weekly new confirmed cases per million population, reaching approximately 60.000, whereas Cambodia had the lowest number at around 250. Excluding Laos, Cambodia also recorded the lowest number of new confirmed deaths, slightly above 10, while Malaysia had the highest value, almost 80. After the peak period, there was a decrease in the number of new confirmed cases and deaths, with a general downward trend observed from January 2023 onwards.

When examining the uniqueness of the COVID-19 situation within each cluster, Cluster 2 in Figure 4(b) stands out with relatively small maximum numbers of weekly new confirmed cases. However, it is important to note that each country in this cluster experienced a significant surge in new confirmed cases prior to the simultaneous peak observed across ASEAN countries. Additionally, despite a relatively high number of new confirmed cases between July 2021 and July 2022, Laos reported comparatively fewer new confirmed deaths during that period. On the other hand, the remaining three clusters witnessed significant peaks in new deaths, albeit occurring after the common peak. Cluster 1 and 3 demonstrate a noticeable contrast in the peak of new confirmed cases among their members, while both groups exhibit similar trends in the number of new confirmed deaths. Notably, these clusters experienced a proportional number of new confirmed deaths with their new confirmed cases. However, Vietnam and Malaysia managed to decrease the number of new confirmed deaths earlier and flatten the curve sooner after the ASEAN new cases' common peak. Singapore in Cluster 4 exhibited a similar fluctuation curve with Brunei throughout the period but reported fewer new confirmed cases and deaths.

On the other hand, the genetics of virus can also be identified by leveraging Biopython's built-in clustering functions to cluster the DNA sequences into four groups based on their country of origin. Table 1 displays that the results were refined and accounted into two or three clusters based on the relatively higher number of sequences appearing in each cluster. The first group consisted of viruses that exactly spread in Brunei, the second group consisted of viruses that may spread in Indonesia and Malaysia, and the third group consisted of viruses that spread in the remaining 7 ASEAN countries.

Table 1. The number of DNA sequences in each cluster based on their country of origin

Countries Name	The number of sequences in			
	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Brunei	9	0	1	0
Cambodia	0	2	8	0
Indonesia	0	5	5	0
Laos	0	0	8	0
Malaysia	0	4	5	1
Myanmar	0	2	7	1
Philippines	0	1	7	2
Singapore	0	0	9	1
Thailand	0	0	9	1
Vietnam	0	1	9	0

As shown in Table 1, 9 out of 10 DNA sequences of the virus that spread in Brunei belong to Cluster 1, which does not contain any sequences from other ASEAN countries. The DNA sequences of the Omicron variant in Indonesia and Malaysia are equally distributed between Cluster 2 and Cluster 3. The DNA sequences from the viruses originating in the remaining seven countries are all found mainly in Cluster 3. Consequently, there are fairly no members in Cluster 4 because the number of sequences contributing to this cluster is relatively small. Overall, compared to clustering ASEAN countries based on their profiles, there appears to be no relationship between the profiles and the spread of viral genome sequences within the countries.

Additionally, using a relatively small dataset, this clustering result is in line with the previous work on genome sequence classification based on country [13], where Brunei's sequence is highly differentiated compared to other ASEAN members. It is interesting to explore why this happens, given that the virus variant in the previous study was Delta, while in this study, we used a later variant, which is Omicron, yet the result for Brunei remains the same. This could be an avenue for future research.

3.2 Comparing COVID-19 situation between ASEAN countries and comparable nation

Involving 159 countries (including ASEAN members) in the dataset, the three most similar profile countries for each ASEAN country along with their similarity indices are listed in Table 2. To illustrate this similarity visually, Figure 5 displays the standardized profile values for Brunei and its comparable countries

in a line graph.

Table 2. The number of DNA sequences in each cluster based on their country of origin

ASEAN Countries Name	The similar profile countries (similarity index)		
Brunei	Kuwait (1.42)	Saudi Arabia (3.18)	United Arab Emirate (3.37)
Cambodia	Djibouti (0.75)	Myanmar (0.81)	Nepal (0.85)
Indonesia	The Philippines (1.41)	Pakistan (3.05)	Vietnam (3.09)
Laos	Cambodia (0.88)	Ghana (1)	Madagascar (1.2)
Malaysia	Bahamas (1.78)	Saudi Arabia (1.98)	Oman (2.09)
Myanmar	Cambodia (0.81)	Djibouti (0.95)	Nepal (1.44)
The Philippines	Indonesia (1.41)	Morocco (1.75)	Nepal (1.99)
Singapore	Bahrain (98.43)	Malta (106.21)	Netherlands (1.34)
Thailand	Chile (1.13)	Costa Rica (1.57)	Colombia (1.64)
Vietnam	Algeria (1.01)	Tunisia (1.35)	Dominican Republic (1.48)

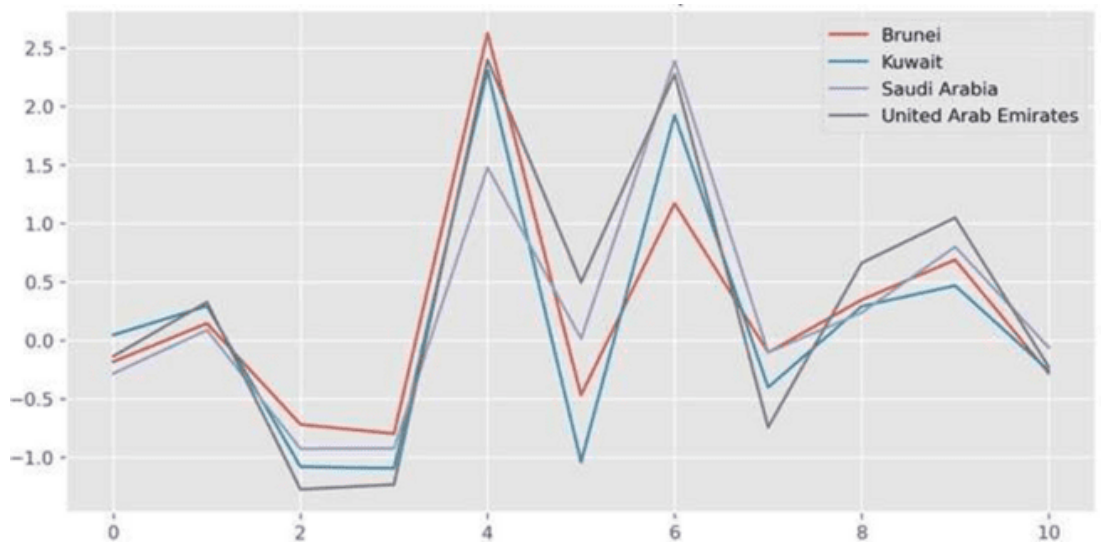
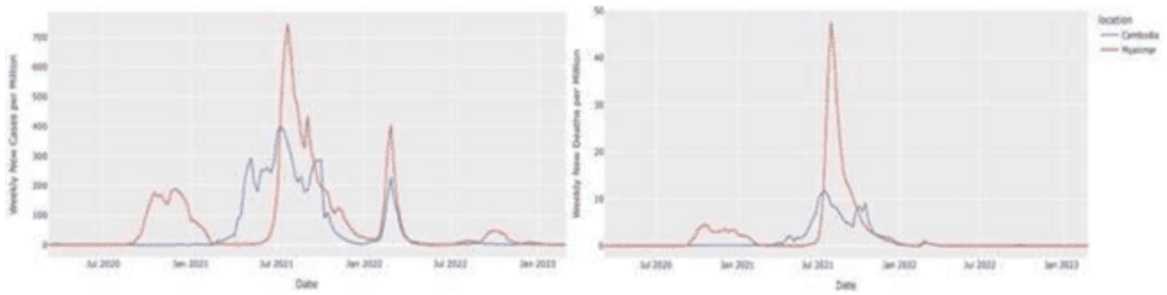


Figure 5. Standardized value of countries' profile that similar to Brunei's profile

From Table 2, it is evident that 5 out of 10 ASEAN countries share a similar profile with countries within the ASEAN region, as indicated by relatively low similarity indices (under 1.5). Interestingly, these five countries were previously grouped in Cluster 2. The COVID-19 situation in these countries can be observed through three combinations of country comparisons (Figure 6): Cambodia-Myanmar, Laos- Cambodia, and Indonesia-Philippines. On the other hand, Singapore, which previously stood in Cluster 4, faces difficulty in finding comparable countries based on its higher similarity index compared to others. For Singapore and the remaining four countries, a five combination of country comparisons is illustrated in Figure 7, comprising Brunei-Kuwait, Malaysia-Bahamas, Thailand-Chile, Vietnam-Algeria, and Singapore-Bahrain.



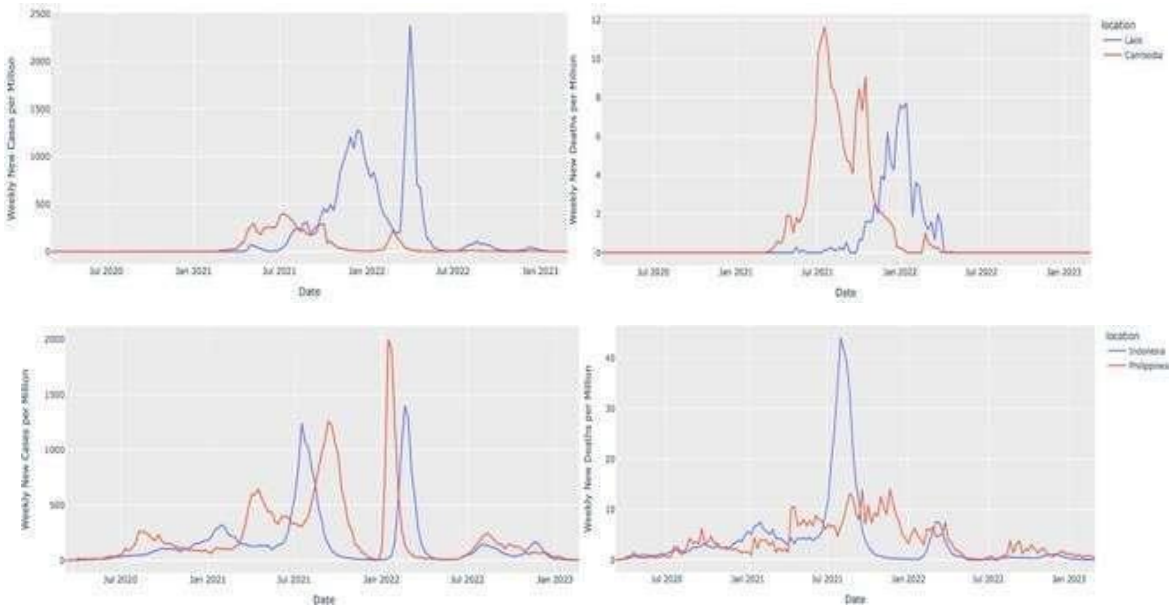
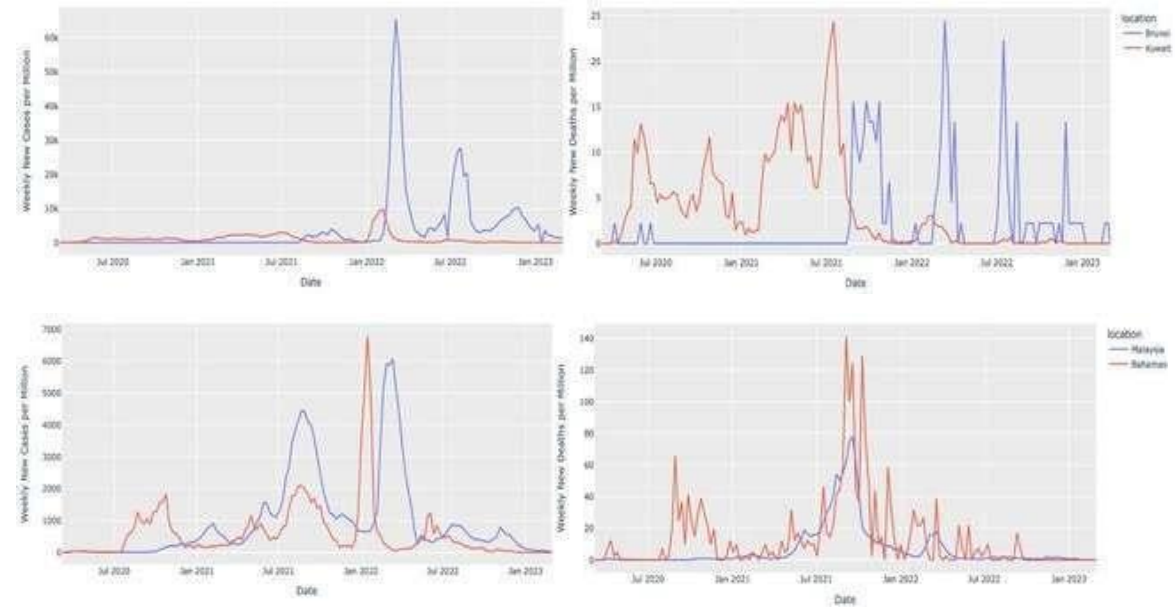


Figure 6. COVID-19 situation between ASEAN countries and comparable nations within the ASEAN region

Analog to cluster observation in the previous subsection, Figure 6 illustrates the similar COVID-19 situations between pairs of countries, particularly at the start, peak, and end of the pandemic. On the other hand, when comparing COVID-19 situations between five ASEAN countries and their comparable nations outside the ASEAN region, Figure 7 illustrates relatively distinct patterns for each pair, particularly regarding the timing of the spread. Firstly, Brunei and Kuwait exhibit different significant fluctuations in the space of new deaths. Interestingly, despite having a fairly equal number of new cases initially, both countries significantly diverge after March 2022, with Kuwait successfully flattening both curves about one year earlier than Brunei. Secondly, Malaysia and the Bahamas show a similar trend in new confirmed cases and deaths from May to December 2021, with Malaysia being relatively successful in unfluctuating the number of new deaths compared to the Bahamas. Thirdly, Thailand and Chile display distinctly different trends in their COVID-19 situations, but both countries can flatten the number of new confirmed deaths after July 2022 with moderately equal numbers. Fourthly, Vietnam and Algeria exhibit a similar pattern in the number of new deaths after July 2021. Lastly, despite having a remarkably high similarity index, Singapore and Bahrain show a similar trend in new confirmed cases and deaths after January 2022, even though both countries experienced a different pattern in their COVID-19 situations prior to that period.



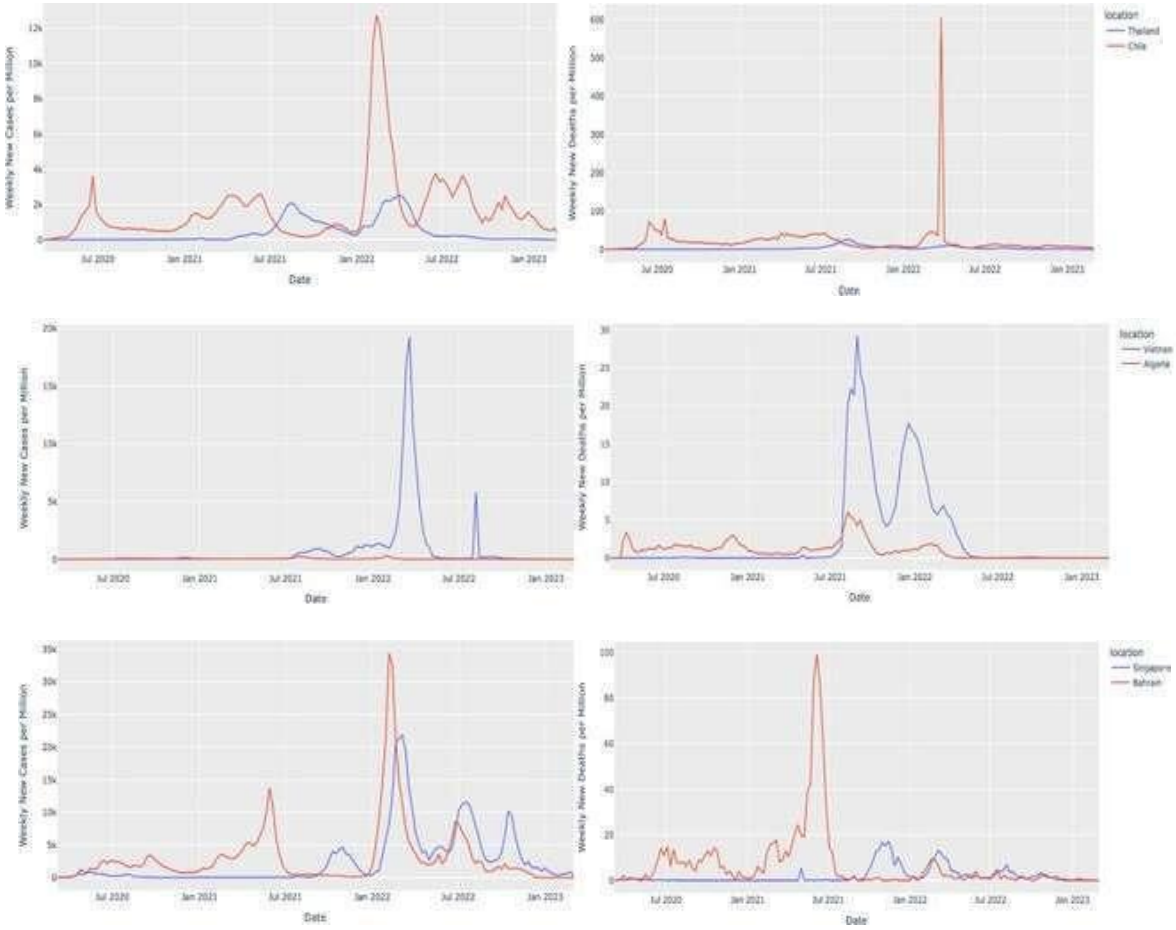


Figure 7. COVID-19 situation between five ASEAN countries and their comparable nations outside the ASEAN region

From the comparison of the COVID-19 situation between ASEAN countries and their comparable nation outside the ASEAN region, it can be inferred that the trend of new cases and deaths is not always the same, even though these countries have similar COVID-19-related profiles. However, when the comparable nation is also within the ASEAN region, the COVID-19 situation pattern is relatively alike. This finding indicates that proximity within the region is an important factor for having a similar pandemic condition besides similar profiles. Nevertheless, comparative studies with other countries with similar profiles can provide valuable insights for effectively managing future pandemics by leveraging analogous resources.

4. Conclusion

This is the first study to observe the COVID-19 situation in 10 ASEAN countries and their comparable nations over three years. Our findings suggest that countries with similar demographic profiles and geographic proximity tend to experience similar pandemic patterns. However, there appears to be no relationship between the spread of identical viral genome sequences among countries and the countries' demographic profiles. Additionally, our genome sequence clustering results, despite using a relatively small dataset, align with previous research on genome sequence classification by country, where Brunei's sequence remains highly differentiated from other ASEAN members. This consistency across studies raises an interesting question for future research on the underlying factors driving this distinction.

By utilizing analogous resources, valuable insights from comparable nations could be useful for effectively managing future pandemics. It is important to note, however, that this study focused solely on the number of weekly new confirmed cases and deaths, which may not provide a comprehensive view of COVID-19. Future studies could address this limitation to enhance efforts in combating the pandemic by drawing more meaningful insights from countries with similar profiles.

References

[1] M. Yüce, E. Filiztekin, and K. G. Özkaya, "COVID-19 diagnosis —A review of current methods," *Biosens Bioelectron*, vol. 172, 2021, doi: 10.1016/j.bios.2020.112752.

[2] Y. H. Jin *et al.*, "A rapid advice guideline for the diagnosis and treatment of 2019 novel coronavirus (2019-nCoV) infected pneumonia (standard version)," 2020. doi: 10.1186/s40779-020-0233-6.

[3] A. Bulgiba, "Adaptive Leadership and Independent Opinions: Lessons From the National COVID-19 Immunisation Programme," *Asia Pac J Public Health*, vol. 34, no. 2–3, 2022, doi: 10.1177/10105395211065309.

[4] N. P. Mishra *et al.*, "Global impacts of pre- and post-COVID-19 pandemic: Focus on socio- economic consequences," 2020. doi: 10.1016/j.sintl.2020.100042.

[5] A. L. Samy, S. Awang Bono, S. L. Tan, and W. Y. Low, "Mental Health and COVID-19: Policies, Guidelines, and Initiatives from the Asia-Pacific Region," 2021. doi: 10.1177/10105395211025901.

[6] R. C. Khanna, M. V. Cicinelli, S. S. Gilbert, S. G. Honavar, and G. V. S. Murthy, "COVID-19 pandemic: Lessons learned and future directions," 2020. doi: 10.4103/ijo.IJO_843_20.

[7] P. Pasquariello and S. Stranges, "Excess mortality from COVID-19: a commentary on the Italian experience," 2020. doi: 10.1007/s00038-020-01399-y.

[8] A. Sanyaolu *et al.*, "Global Pandemicity of COVID-19: Situation Report as of June 9, 2020," *Infectious Diseases: Research and Treatment*, vol. 14, 2021, doi: 10.1177/1178633721991260.

[9] H. M. Baniamin, M. Rahman, and M. T. Hasan, "The COVID-19 Pandemic: Why are Some Countries More Successful than Others?," *SSRN Electronic Journal*, 2020, doi: 10.2139/ssrn.3575251.

[10] Max Roser, Hannah Ritchie, Esteban Ortiz-Ospina, and Joe Hasell, "Coronavirus Disease (COVID-19) – Statistics and Research - Our World in Data," *University of Oxford*, vol. 2020, 2020.

[11] S. Rampal *et al.*, "The epidemiology of covid-19 in ten southeast asian countries," *Medical Journal of Malaysia*, vol. 76, no. 6, 2021.

[12] GISAID, "GISAID Initiative," 2020.

[13] E. Khatizah and H. S. Park, "Country-Based COVID-19 DNA Sequence Classification in Relation with International Travel Policy," *Applied Sciences (Switzerland)*, vol. 14, no. 5, 2024, doi: 10.3390/app14051916.

.