SEMARAK ILMU PUBLISHING

Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

Calendar Heatmap Visualization for Descriptive Analytics of Epidemics: A Case Study of Covid-19

Maslina Zolkepli[1], Isa Ramlan[1,*]

[1] Department of Computer Science, Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia

ARTICLE INFO

Article history:
Received 26 June 2023
Received in revised form 24 July 2023
Accepted 5 October 2023
Available online 2 December 2023

Keywords:
Covid-19; descriptive analytics; calendar heatmap; D3 algorithm

ABSTRACT

The Covid-19 epidemic has presented unparalleled obstacles to the world, including Malaysia and Indonesia. A visual analytics approach known as the Covid-19 calendar heatmap visualization is proposed to provide relevant and timely information about Covid-19 daily cases and deaths in Malaysia. The aim of this approach is to offer a platform for various stakeholders to make informed decision based on descriptive analytics. The Covid-19 information visualization approach allows users to compare different types of visualizations such as calendar heatmaps, line charts, and bar charts to detect Covid-19 trends or patterns. The approach is implemented as a web-based system using D3 JavaScript library and tested using the data collected from Coronavirus Pandemic Data Explorer. Results show that the Covid-19 calendar heatmap visualization approach is an invaluable approach for understanding epidemic trends as it provides users with easy access to visualizations and fast understanding of Covid-19 situations such as daily cases, deaths and recoveries. The target users for the proposed approach are the National Security Council and Ministry of Health staff, company crews, and the general public. Future upgrades to the approach includes adding additional attributes such as the number of tests and recoveries to assist users in making predictive analytics using machine learning models.

1. Introduction

In recent times, many countries in the world witnessed the beginning of the end of the Covid-19 epidemic. However, as time progressed, a relatively stable situation unfolded. The emerging variants of the coronavirus observed thus far have shown a close resemblance to Omicron, with no significant alterations in its overall impact [1] and Covid-19 remains a global health threat and experts continue to monitor its spread closely [2].

In Malaysia, the Ministry of Health is maintaining its implementation of Covid-19 control and prevention measures according to established protocols [3] as Covid-19 cases continues to emerge and it is posing a significant risk of death, especially for elderly individuals with underlying health conditions. Therefore, it is crucial to conduct an in-depth analysis of the factors contributing to the

* *Corresponding author.*
E-mail address: isaramlan@gmail.com

https://doi.org/10.37934/araset.34.2.117

country's struggles with controlling the spread of the virus [4]. Prolonged periods of lockdown and social distancing measures have also had severe economic and social consequences, making it imperative to find effective solutions to combat the epidemic.

Data analytics plays a crucial role in understanding and addressing the complexities of the Covid-19 epidemic [5]. In the battle against the Covid-19 epidemic, data analytics emerges as a powerful tool, providing valuable insights and aiding decision-making processes. At the outset of the epidemic, users were compelled to manually search for Covid-19 cases [6], which was akin to searching for a document manually. The process required visualizing data by examining individual data points one-by-one, which can have detrimental effects on the brain and eyes due to prolonged screen exposure. Furthermore, the manual search and visualization process is time-consuming and dependent on individual imagination, resulting in a less efficient use of resources.

In this research, a Covid-19 descriptive visualization approach through calendar heatmap [7] is proposed where it can provide users with an intuitive and comprehensive overview, facilitating easy analysis of the epidemic's impact. The insights acquired by the calendar heatmap visualization is targeted to contribute to the on-going efforts to combat the epidemic and minimize its economic and social consequences. The visualization approach also enables users to discern patterns in the data visualization. The result will show the differences between the primary visualization methods used, which is the Calendar Heatmap, and other visualization techniques such as the Bar Chart and Line Chart.

There are a variety of Covid-19 visualization approaches [8-10] introduced by companies, universities, hospitals and researchers. Our World in Data [11] produces data visualization process the world problem Issues. It makes the knowledge on big problems accessible and understandable. The data visualization can be changed based on what the users need to see. The visualization of Our World in Data is shown in Figure 1. Its Covid-19 data visualization and analysis include a range of metrics such as case counts, testing rates, mortality rates, and vaccination rates. The website allows users to explore the data through interactive visualizations, tables, graphs, and heatmaps. Our World in Data is well-regarded for its data quality, methodology, and transparency in sourcing and presenting data.

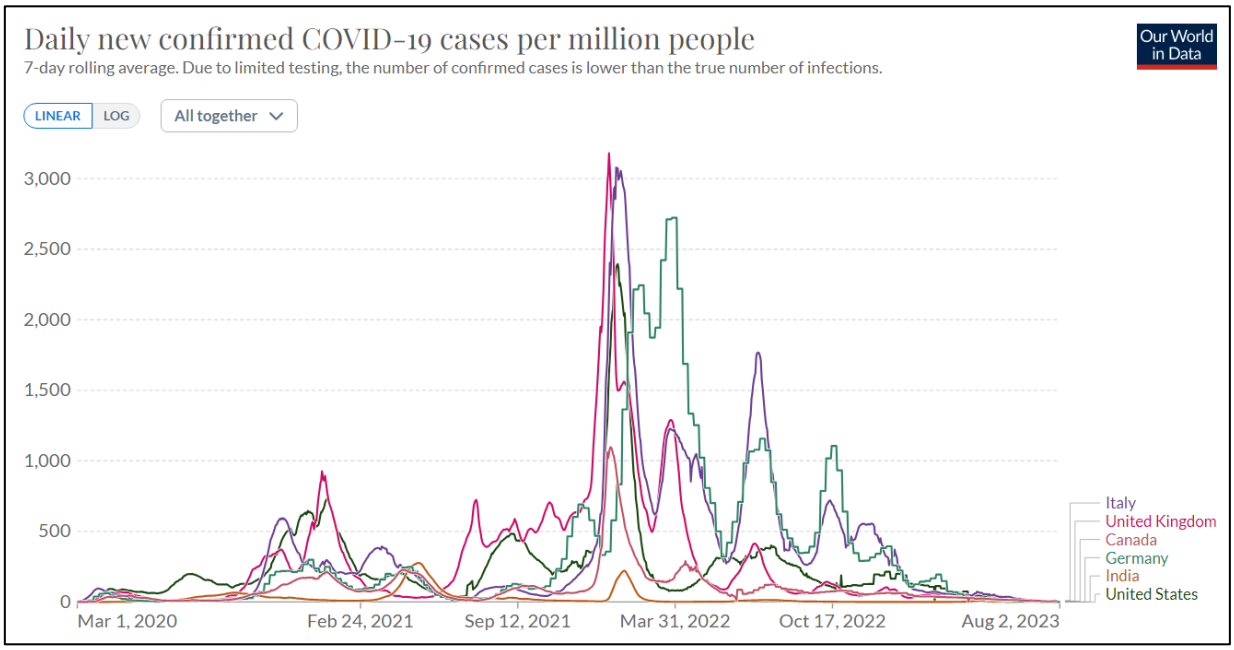

Fig. 1. The main interface of Our World in Data website

Worldometer [12] is a real-time statistics website that provides various information, including the Covid-19 epidemic. It has been widely used by researchers and policymakers. The website updates Covid-19 data from multiple sources, including the World Health Organization (WHO) and national health agencies. Worldometer provides data on confirmed cases, deaths, and recoveries, as well as graphs and maps to visualize the data. The website also provides demo-graphic breakdowns of the data, such as age, gender, and pre-existing conditions. Worldometer's Covid-19 data has been cited in various research studies and media reports. Worldometer visualization is shown in Figure 2.

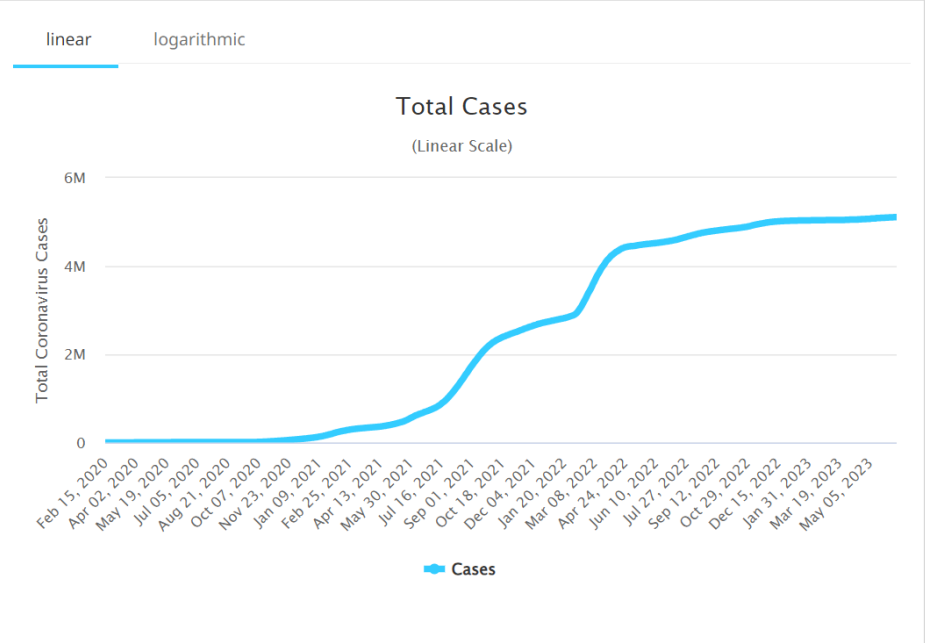

Fig. 2. Worldometer Covid-19 cases visualization

The Covid-19 Tracker by the New York Times [13] is a source for tracking the Covid-19 epidemic in the United States. The visualization of the Covid-19 Tracker by the New York Times is shown in Figure 3. It provides up-to-date information on the number of cases, deaths, and hospitalizations in each state and county, as well as trends over time using line charts. It includes informative visualizations to help users better understand the spread of the virus. The New York Times uses data from various sources, including state and local health departments, to provide accurate and reliable information to its readers. Overall, the Covid-19 Tracker by the New York Times is targeted to individuals, healthcare professionals, and policymakers alike, in their efforts to combat the Covid-19 epidemic.

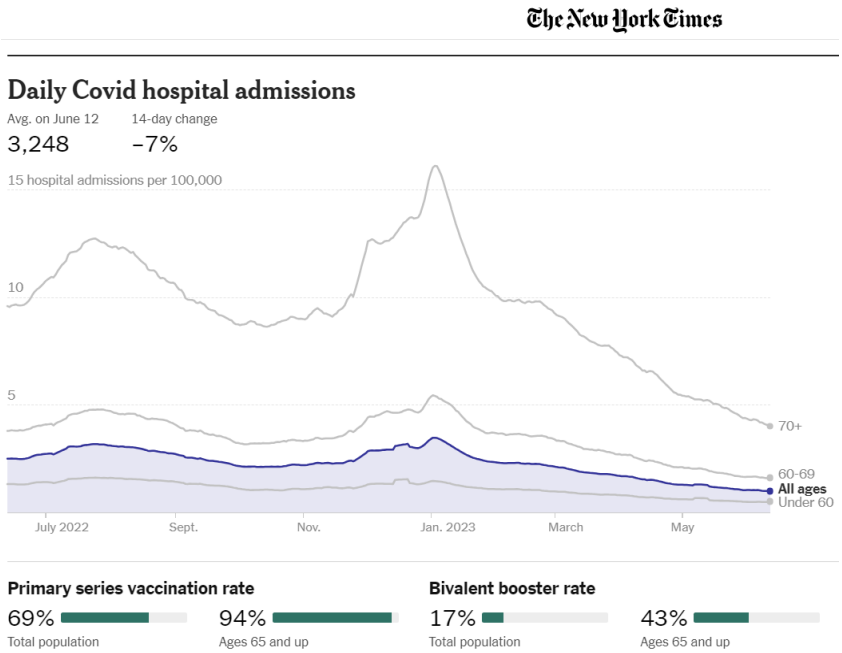

Fig. 3. The Covid-19 Tracker by the New York Times

The Covid-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University [14] is a tool for tracking the global spread of the Covid-19 epidemic, as shown in Figure 4. The dashboard provides up-to-date information on confirmed cases, deaths, and recoveries, broken down by country, region, and state. The data is presented in a variety of visualizations, including maps, charts, and tables, allowing users to easily track trends and patterns in the spread of the virus.

The dashboard has been used by public health officials and policymakers as a key source of information in their decision-making processes. It has also been used by the media to report on the latest developments in the epidemic.

Overall, the Covid-19 Dashboard by the CSSE at Johns Hopkins University has emerged as a valuable resource in the fight against the Covid-19 epidemic, providing critical information and insights to help policymakers, public health officials, and the public at large better understand and respond to the ongoing crisis.

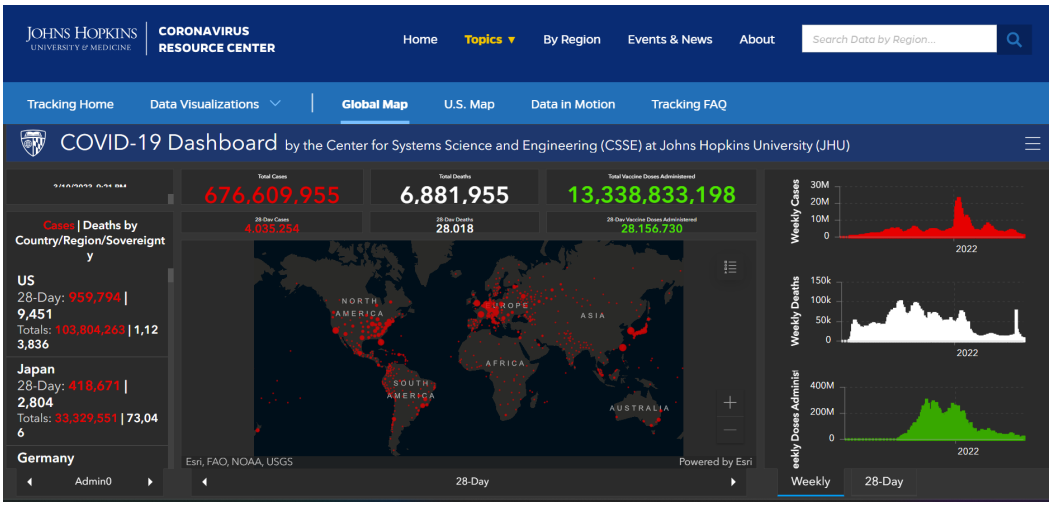

Fig. 4. Covid-19 Dashboard at Johns Hopkins University

Ministry of Health Malaysia also offers visualization and analytics of Covid-19 trends [15,16] as shown in Figure 5. It displayed many kinds of information about Malaysians trending health issues. The main page will display on data visualization about trending disease. It also holds campaigns for Malaysian citizen to fight the trending disease.

Fig. 5. The Ministry of Health Malaysia Website

The World Health Organization's Covid-19 visualization [17,18] is a data-driven tool provided by the World Health Organization to effectively present and analyze Covid-19 related information. This visualization platform offers interactive charts, maps, and graphs that allow users to explore and understand various aspects of the epidemic. It provides bar chart representations of key data such as the number of confirmed cases, deaths, and recoveries globally and across different countries. The visualization tool also allows users to examine trends over time, compare data between regions, and assess the impact of interventions and measures implemented to control the spread of the virus. The World Health Organization's Covid-19 visualization in shown in Figure 6.

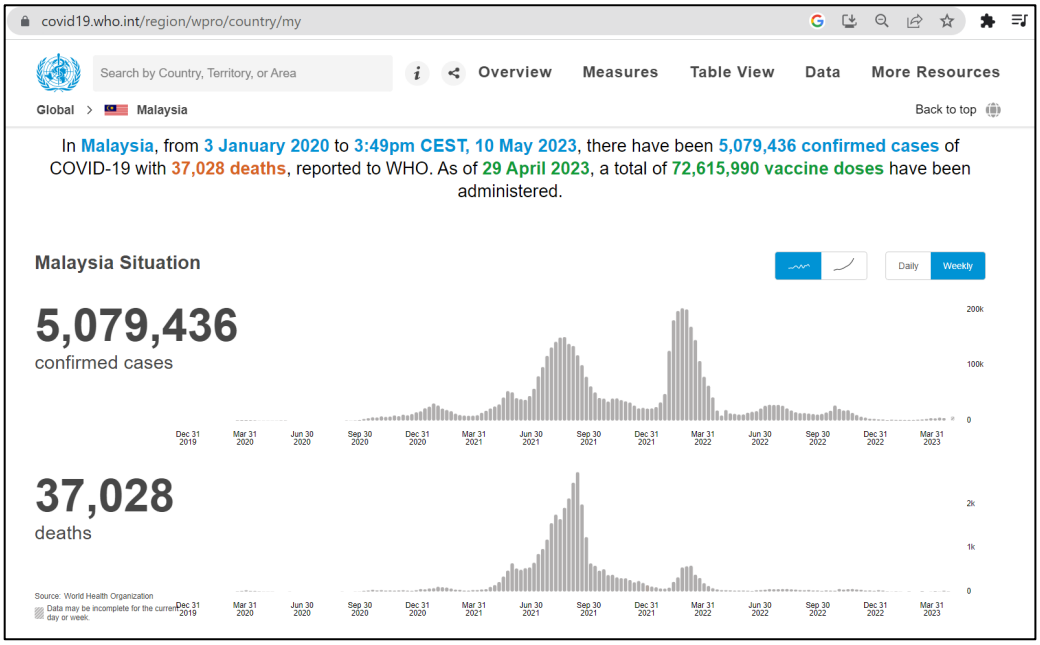

Fig. 6. The World Health Organization Covid-19 Visualization

All of the above Covid-19 visualization approaches vary in their focus and offerings. Our World in Data and the New York Times Covid-19 Tracker emphasize data analysis and visualizations, Worldometer provides real-time statistics, the Johns Hopkins Covid-19 Dashboard is renowned for accuracy, Ministry of Health Malaysia website offer localized information, and the WHO Covid-19 Dashboard provides global data from a trusted source. Users can choose the approach that aligns with their specific needs, whether it's accessing global trends, tracking real-time updates, or finding localized information. The comparison of Covid-19 visualizations are shown in Table 1.

Table 1
Comparison of Covid-19 Visualization Websites

Covid-19 Visualization Websites	Target User	Type of Visualization
Our World in Data	General Public	Line Graph
Worldometer	General Public	Bar Graph, Line Graph
New York Times Covid-19 Tracker	General Public	Line Graph
Johns Hopkins Covid-19 Dashboard	General Public	Maps, Bar Graph
Ministry of Health Malaysia Website	Malaysian citizens	Text, Numbers
WHO Covid-19 Dashboard	General Public	Maps, Bar Graph

Calendar heatmap visualization is a powerful tool for descriptive analytics of epidemics. It can be used to identify hot spots of infection that can help public health officials to target their resources and interventions to the areas where they are most needed. Calendar heatmap is able to track the spread of the epidemic over time and patterns in the spread of the disease can be identified and predictions of future spread of the disease can be made. Calendar heatmap can also be used to compare different epidemics. It helps to identify factors that contribute to the spread of epidemics and to develop more effective prevention and control strategies. From the epidemic data point of view, calendar heatmaps are particularly useful for visualizing epidemic data because they are able to show multiple variables at the same time. It helps to identify relationships between different variables, such as the relationship between the number of cases and the number of deaths. Trends in the epidemic data is also able to be highlighted, where it can help to identify changes in the spread of the epidemic over time. Calendar heatmaps also makes it easier for people to understand the data and to identify patterns.

Heatmaps are used to track the spread of the Ebola virus in West Africa [19]. The calendar heatmaps showed that the virus was spreading rapidly in some areas, and this information helped to guide the deployment of resources and interventions. Heatmaps are also beneficial in comparing the spread of the Zika virus in Brazil and Colombia [20]. The heatmaps showed that the virus was spreading more rapidly in Brazil, and this information helped to explain why Brazil was experiencing a larger outbreak. In the most recent pandemic, calendar heatmaps are used to track the spread of the COVID-19 pandemic around the world. The calendar heatmaps showed that the pandemic was spreading rapidly in some countries, and this information helped to guide the global response to the pandemic. As the use of calendar heatmaps becomes more widespread, it is expected to see even more innovative ways to use this powerful tool to understand and respond to epidemics.

The remaining sections of this article are structured as follows. Section 2 presents the methodology implemented in this research. The implementation of the research is discussed in detail, particularly the descriptive analytics approach, the Covid-19 dataset, the visualization framework, and the calendar heatmap visualization approach using D3.js algorithm. Section 3 focuses on the results of the experiment and analysis. Finally, Section 4 provides a conclusion, summarizing the content derived from the research outcomes. Suggestions and future work to improve this research are also highlighted.

2. Methodology

2.1 Data Analytics Life Cycle for Covid-19 Descriptive Analytics Research

The Data Analytics Life Cycle [21,22] for Covid-19 is a structured approach that encompasses various stages to extract meaningful insights from data related to the Covid-19 epidemic. It involves processes such as data collection, data cleaning and preprocessing, exploratory data analysis, and the application of statistical and visualization techniques to describe and summarize the Covid-19 data. By following this life cycle, researchers can uncover patterns, trends, and correlations in the data, leading to a comprehensive understanding of the descriptive aspects of the epidemic. The data analytics life cycle for Covid-19 is shown in Figure 7.

In the early days of the COVID-19 pandemic, descriptive data visualization was used to track the spread of the virus and to identify hot spots. As the pandemic has progressed, descriptive data visualization has also been used to monitor the effectiveness of interventions, to track the progress of the pandemic, and to communicate the risks of the virus to the public.

Descriptive data visualization is a powerful tool that can be used to understand and communicate the spread of COVID-19. It can be used to visualize a variety of data, including the number of cases, the number of deaths, the age distribution of cases, and the geographic distribution of cases. Some of the most common descriptive data visualization approaches used to visualize COVID-19 data include bar charts, line charts, pie charts, map visualizations, and many more. Bar charts are a simple and effective way to visualize the number of cases or deaths. It can be used to compare the number of cases or deaths over time, or to compare the number of cases or deaths between different countries or regions. Line charts are another effective way to visualize the number of cases or deaths over time. They can be used to show trends in the spread of the virus, or to identify spikes in the number of cases or deaths. Pie charts can be used to visualize the distribution of cases or deaths by a particular variable, such as age or gender. They can be helpful for understanding who is most at risk from the virus. Map visualizations can be used to visualize the geographic distribution of cases or deaths. They can be helpful for understanding where the virus is spreading, and for identifying areas that are most affected by the epidemic. These are just a few of the many descriptive data visualization approaches that can be used to visualize COVID-19 data. The best approach to use will depend on the specific data that is being visualized, and the target audience.

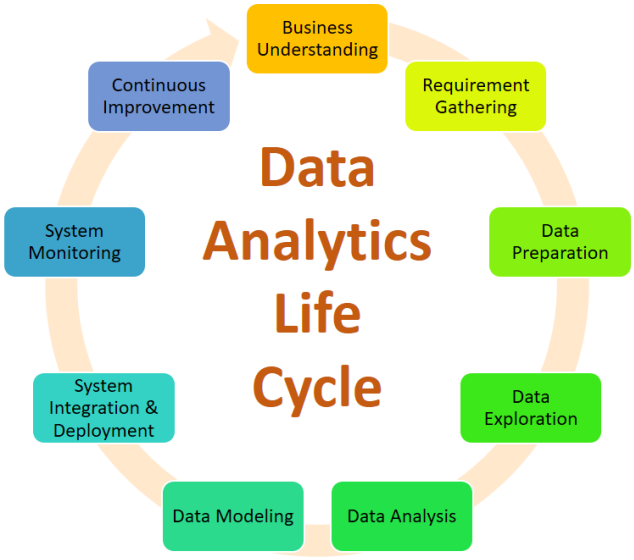

Fig. 7. Data Analytics Life Cycle

2.2 Covid-19 Dataset

The dataset is sourced from Our World in Data Websites [23], specifically the Coronavirus Pandemic Data Explorer. The dataset encompasses various attributes as outlined in Table 2.

Table 2
Covid-19 Dataset in Malaysia

Data Attribute	Data Description
Location_ID	Unique location ID
Location	Name of location
Date	Date of event
New cases	Number of new cases daily
Total cases	Number of accumulative cases
New Death	Number of new deaths daily
Total Death	Number of accumulative deaths

2.3 Covid-19 Descriptive Analytics and Visualization Framework

The Covid-19 Descriptive Analytics and Visualization framework refers to the structure that enables the analysis and visualization of data related to the Covid-19 epidemic, as shown in Figure 8. The framework incorporates descriptive analytics and visualization tools to extract meaningful insights and present them in a visually appealing and understandable manner. By employing this framework, researchers, policymakers, and the general public can gain valuable insights into the spread, impact, and trends of the Covid-19 virus, aiding in decision-making and resource allocation to combat the epidemic effectively.

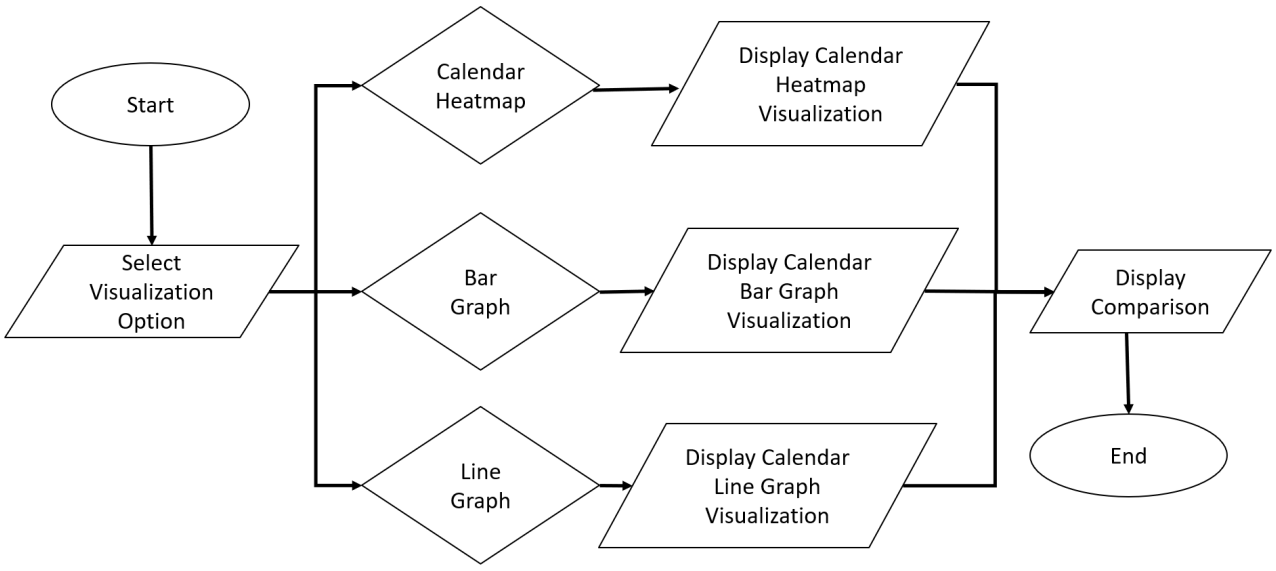

Fig. 8. Covid-19 Data Visualization Framework

2.4 Calendar Heatmap Visualization of Covid-19 outbreak cases using D3.js algorithm

Calendar heatmaps are a powerful visualization technique that can be used to represent time-series data. They are often used in business intelligence and analytics applications, and they are becoming increasingly popular for social media analytics.

A general framework for creating calendar heatmaps has been proposed [24], which discusses the advantages and disadvantages of the technique used with temporal data. A study on the effectiveness of calendar heatmaps for visualizing time-series data [25] found that they can be effective for communicating temporal patterns, but they can be difficult to read if the data is too dense.

There has been a growing interest in the use of calendar heatmaps for social media analytics. A method has been presented for creating calendar heatmaps from Twitter data, which was used to visualize the frequency of tweets about the 2012 US presidential election [26].

The D3.js calendar algorithm [27-29] is a method for creating interactive calendar heatmaps. It is a powerful tool for visualizing time-series data, and it is becoming increasingly popular for social media analytics. The algorithm works by first creating a grid of cells, each representing a day of the year. The cells are then colored according to the value of the data being visualized. For example, if the data represents the number of COVID-19 cases, the cells would be colored red for high values and green for low values.

The D3.js calendar heatmap algorithm also allows users to interact with the heatmap. Users can zoom in and out of the heatmap, and they can also click on individual cells to see more information about the data on that day. This interactivity makes it easier for users to understand the data and to identify patterns. The D3.js calendar algorithm is a versatile tool that can be used to visualize a variety of time-series data. It is easy to use and it is highly customizable, making it a popular choice for both beginners and experienced developers. A pseudocode representation of the D3.js calendar algorithm is shown in Figure 9.

```
function createCalendarHeatmap(data) {
  // Create a grid of cells, each representing a day of the year.
  var cells = d3.range(365).map(function(index) {
    return {
      day: index,
      value: data[index]
    };
  });

  // Color the cells according to the value of the data.
  var colorScale = d3.scaleLinear()
    .domain([0, d3.max(data)])
    .range(["red", "green"]);

  cells.forEach(function(cell) {
    cell.color = colorScale(cell.value);
  });

  // Create a div to hold the heatmap.
  var div = d3.select("body").append("div");
  div.attr("width", "500px");
  div.attr("height", "500px");

  // Create a canvas to draw the heatmap on.
  var canvas = div.append("canvas");
  canvas.attr("width", "500px");
  canvas.attr("height", "500px");

  // Draw the heatmap on the canvas.
  var ctx = canvas.getContext("2d");
  cells.forEach(function(cell) {
    ctx.fillStyle = cell.color;
    ctx.fillRect(cell.day * 10, 0, 10, 10);
  });
}
```

Fig. 9. D3.js calendar heatmap pseudocode

One of the main contributions of calendar heatmaps using the D3 algorithm to the understanding of COVID-19 outbreak cases is their interactivity. Users can zoom in and out of the heatmap, and they can also click on individual dates to see more information about the number of cases on that date. This interactivity makes it easier for users to understand the spread of the virus.

Calendar heatmaps using the D3 algorithm can also be customized. Users can change the colours of the heatmap, the size of the text, and the other visual elements of the heatmap. This customization allows users to create heatmaps that are tailored to their specific needs.

Several organizations have created calendar heatmaps to visualize the spread of COVID-19 cases. The New York Times has created a heatmap that shows the number of COVID-19 cases in the United States by date [13]. The BBC has created a heatmap that shows the number of COVID-19 cases around the world by date [30]. The Johns Hopkins University Center for Systems Science and Engineering has created a heatmap that shows the number of COVID-19 cases in different countries by date [14].

In conclusion, calendar heatmaps are a powerful tool for visualizing COVID-19 outbreak cases. They can help to identify hot spots, track the progression of the pandemic, and make predictions about the future course of the pandemic. This information can be used by public health officials to make informed decisions about how to respond to the pandemic.

3. Calendar Heatmap for Covid-19 Visualization Results

The Covid-19 visualization approaches are implemented on a website that includes three primary types: calendar heatmap, bar chart, and line chart. Users have the flexibility to select the visualization type that best suits their needs and preferences.

Once users select their preferred visualization type, the system promptly generates and displays the chosen data visualization. This interactive feature allows users to engage with the data, explore trends, and derive insights specific to their interests. By presenting data in various visual formats, the website caters to different analytical needs and enhances the users' understanding of the Covid-19 situation. The proposed Covid-19 Visualization Approach main page is shown in Figure 10.

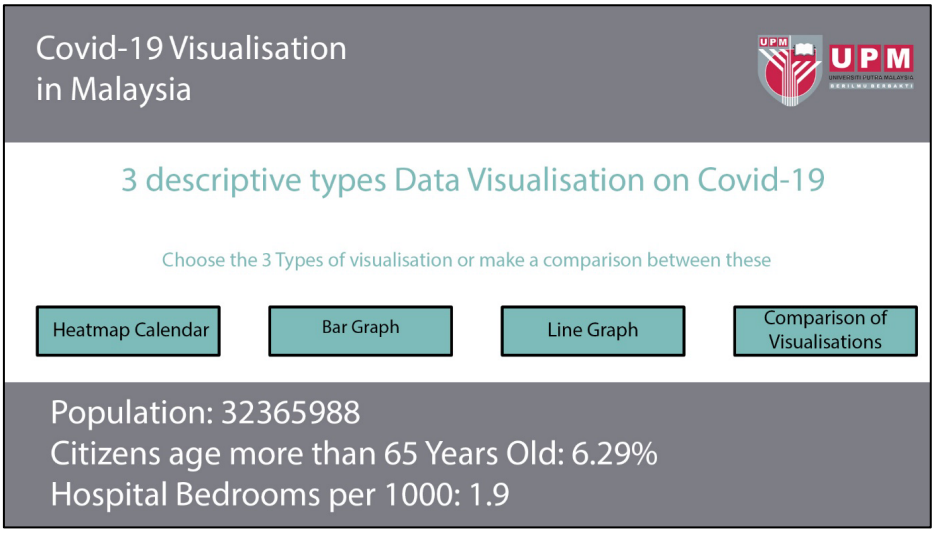

Fig. 10. Covid-19 Visualization Main Page

The calendar heatmap visualization specifically designed to showcase the daily cases and daily death cases related to Covid-19. This visualization provides a unique and comprehensive way for users to analyze and understand the impact of the epidemic over time. The Calendar Heatmap represents each day of a given time period as a separate cell or tile in a calendar-like grid. The cells

are filled with colors that correspond to the intensity or magnitude of the Covid-19 cases or deaths for that particular day. This color-coded representation allows users to discern patterns and variations at a glance, as different colors indicate varying levels of intensity.

The accuracy of the data values displayed in the calendar heatmap is of paramount importance. Users can trust that the data presented aligns with the latest and most reliable information available, ensuring the credibility of the visualization. This accuracy ensures that users can make informed interpretations and draw meaningful insights from the calendar heatmap.

By utilizing different colors to differentiate the intensity, the calendar heatmap effectively highlights the varying degrees of Covid-19 cases or deaths over time. This color differentiation allows users to identify hotspots, areas of concern, or regions with particularly high or low levels of infection or mortality. It helps to identify trends, patterns, and areas requiring focused attention or intervention. The calendar heatmap visualizations of Covid-19 new cases and death cases in Malaysia are shown in Figure 11. Each square in the calendar heatmap represents one day of the year. Days with low indices are represented in light aqua, and as the index goes up, the color changes to aqua, green, yellow, orange, and red. Days with the highest indices are represented in black. Table 3 summarizes the color coding scheme.

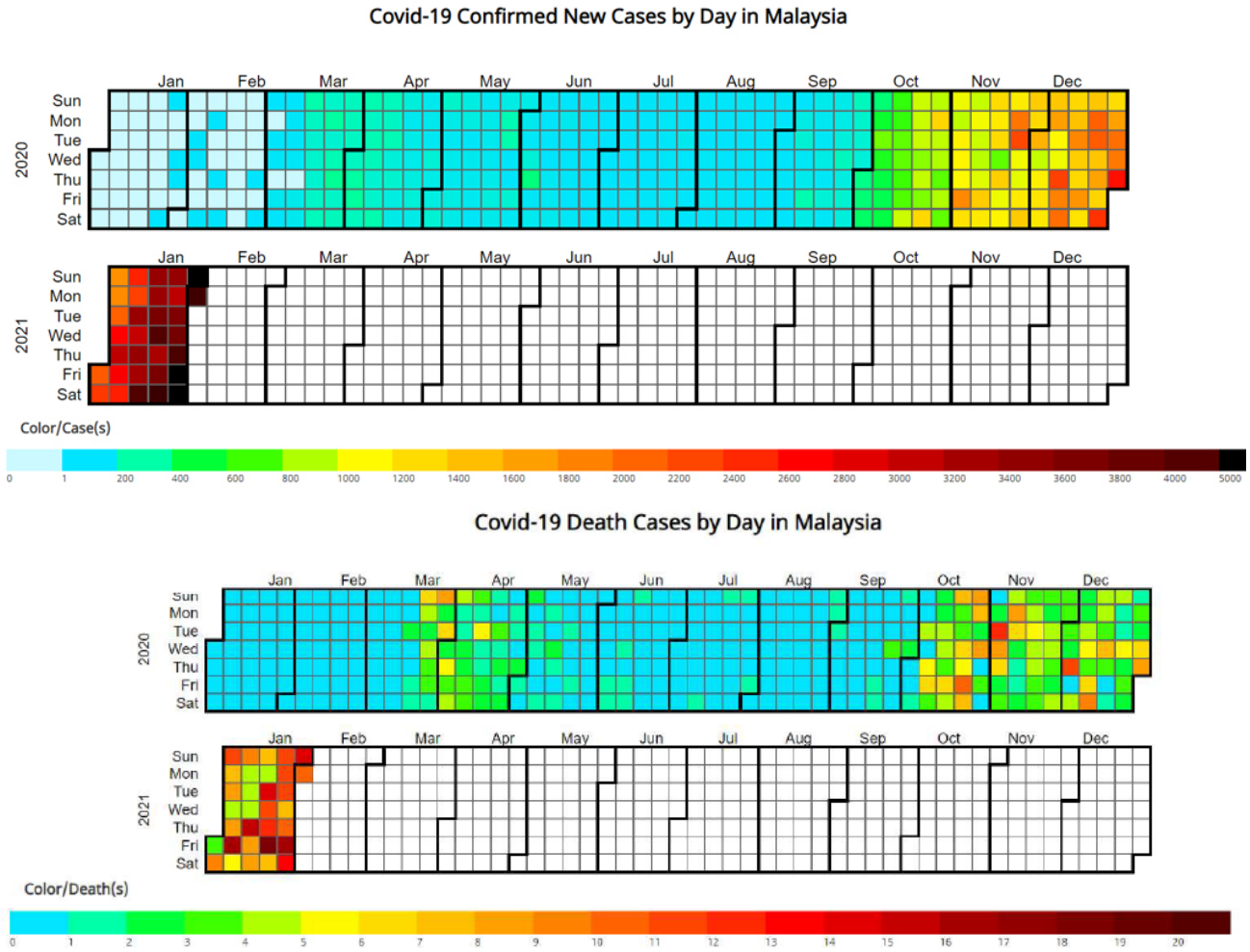

Fig. 11. Calendar Heatmap Visualization for Covid-19 Cases in Malaysia

Table 3
Color coding scheme for calendar heatmap visualization of Covid-19 cases in Malaysia

Index	Color
Very Low	Light aqua
Low	Aqua
Medium	Green
High	Yellow
Very high	Orange
Extremely high	Red
Peak	Black

Overall, the calendar heatmap visualization offers users a visually compelling and accurate representation of daily Covid-19 cases and deaths. It is able to facilitate the understanding of the epidemic's impact, supports data-driven decision-making, and enhances the communication of critical information related to the spread and severity of the virus.

The "Bar Graph" page as shown in Figure 12 presents a visualization of daily cases and daily death cases through a bar graph. Users can accurately visualize the bar graph, where each bar represents the corresponding data value and distinguishes it from others. The "Bar Graph" page offers a visual representation of daily cases and daily death cases using a bar graph. In this visualization, each bar corresponds to a specific data value and provides a clear distinction between different values.

Bar graphs are particularly effective in displaying discrete and categorical data, making them well-suited for representing Covid-19 statistics. Each bar in the graph represents a specific category, such as a specific date or location, and the length or height of the bar corresponds to the magnitude or count of the data value it represents. Users can accurately interpret the bar graph by observing the relative lengths or heights of the bars. This allows for easy comparison between different categories and helps identify patterns, trends, or significant variations in the data. For example, users can quickly identify which days had the highest number of cases or compare the number of deaths across different regions. Furthermore, the visual nature of the bar graph aids in the communication of data to a wide audience. It simplifies complex information, making it more accessible and understandable, even for individuals without a strong statistical or analytical background.

The "Line Graph" page as shown in Figure 13 presents a line graph representing daily cases and daily death cases. Users can visualize the line graph with accurate data values, and the line differentiates the values over time. The "Line Graph" page displays a visual representation of daily cases and daily death cases using a line graph. This type of visualization is particularly effective in illustrating trends and changes over time. In the line graph, each data point is represented by a marker, and these markers are connected by a line, creating a continuous visual representation of the data over a specific time period. The vertical axis typically represents the number of cases or deaths, while the horizontal axis represents the progression of time, such as days, weeks, or months.

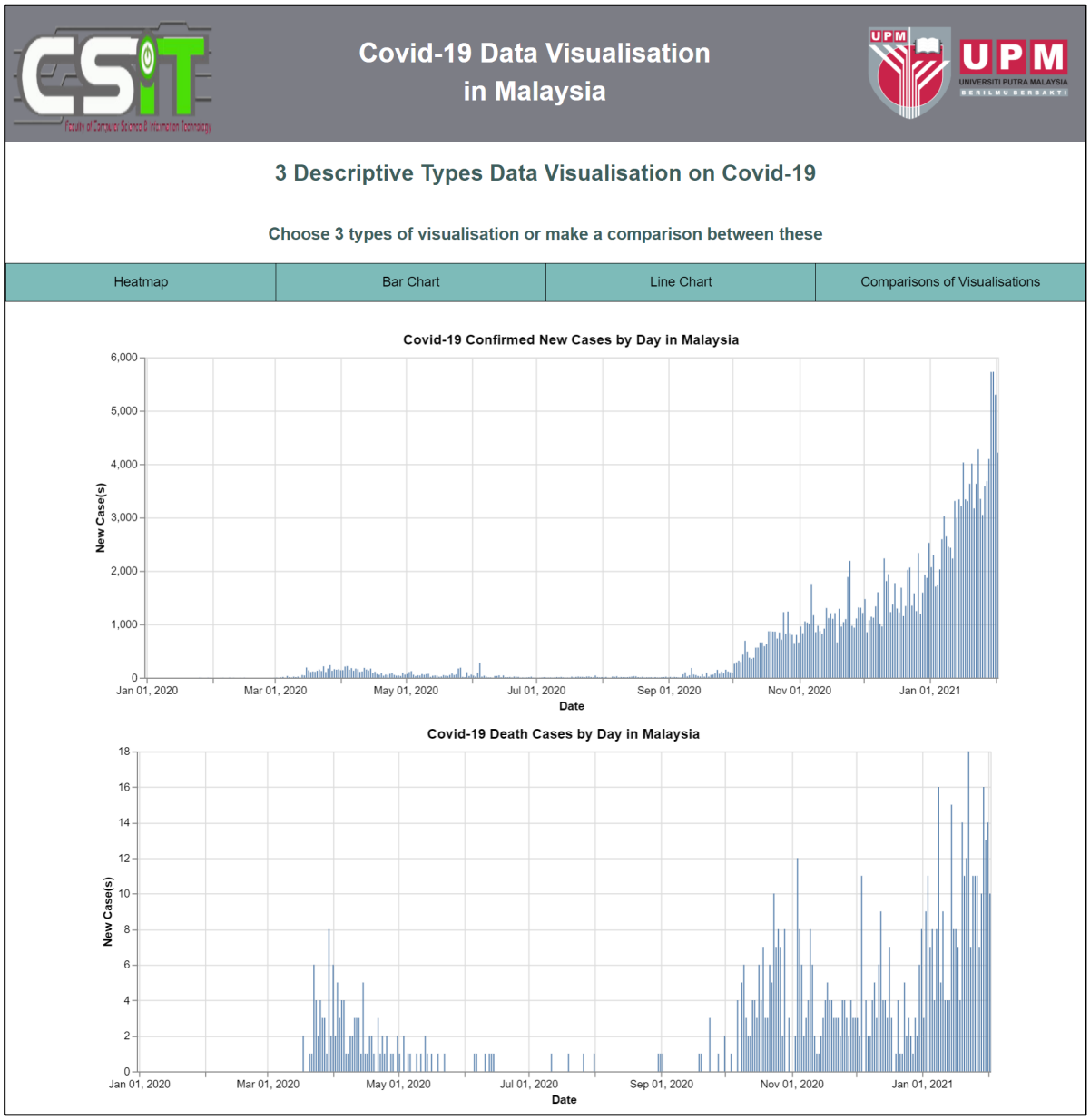

Fig. 12. Result on Bar Graph Visualization

By plotting the data points and connecting them with a line, the line graph provides a clear depiction of the fluctuation and progression of daily cases and deaths over time. This enables users to visualize the rise or fall of the numbers, identify peak periods, and observe any significant patterns or trends. Users can accurately interpret the line graph by examining the position of the line at specific time points. The line graph distinguishes the values by the slope, height, or position of the line relative to the axis. For example, an upward or downward slope indicates an increase or decrease in cases or deaths, respectively.

The line graph's ability to showcase temporal patterns and changes makes it valuable for tracking the progression of the epidemic, assessing the effectiveness of interventions, and identifying potential turning points. By visualizing the data with accuracy, users can gain insights into the impact of the virus over time and make informed decisions based on the observed trends.

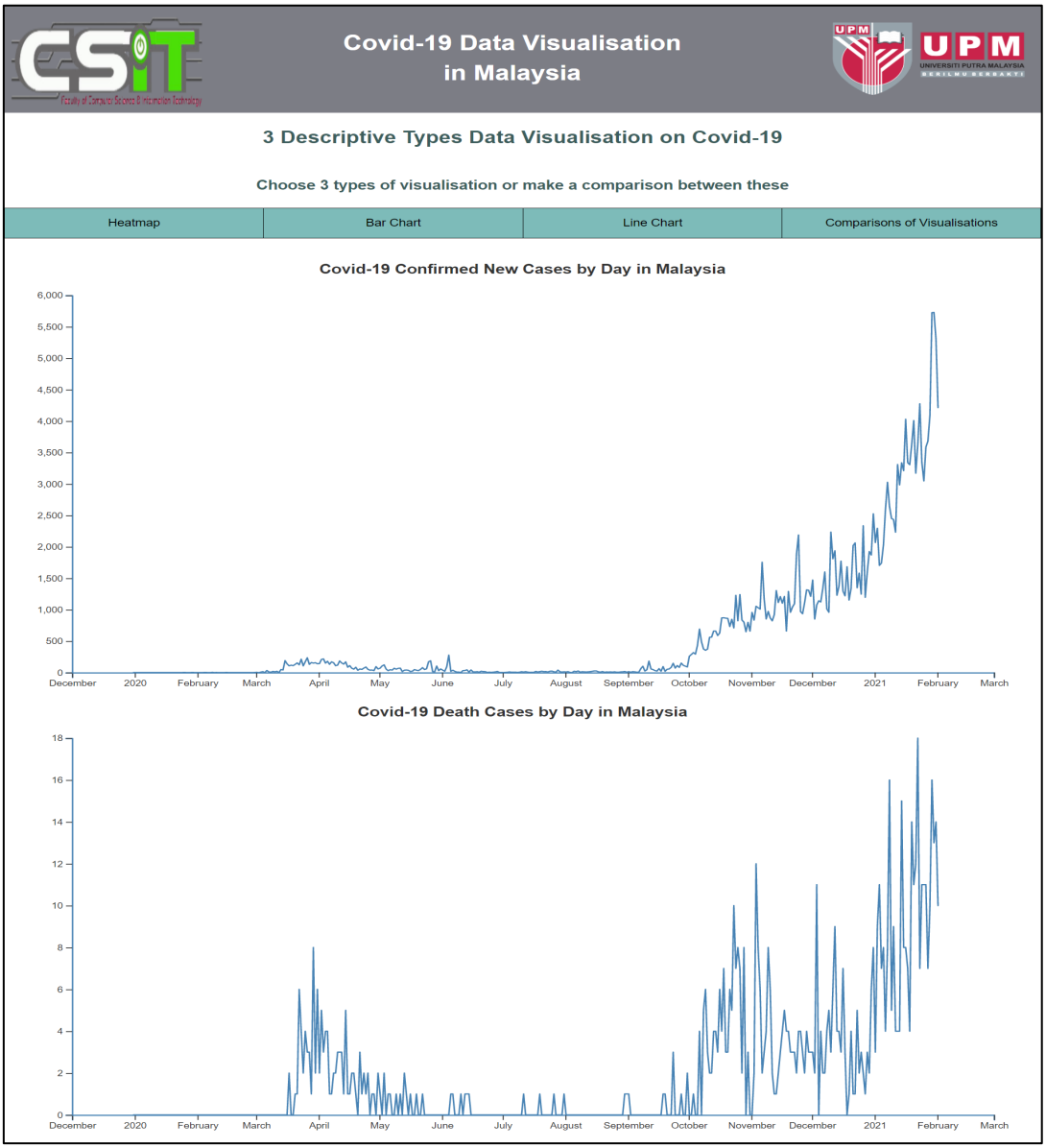

Fig. 13. Result on Line Graph Visualization

3.1 Discussion on the comparison of calendar heatmap, bar chart and line chart visualization of Covid-19 dataset

The calendar heatmap visualization stands out with its ability to provide an intuitive and comprehensive overview of the data. It typically uses colors to represent the intensity or magnitude of a particular variable, allowing users to quickly identify patterns and variations in the data. Calendar heatmaps are especially useful for representing large datasets or time series data, as they condense information into a visual grid format. In contrast, a bar chart is effective in displaying discrete and categorical data. It uses vertical or horizontal bars to represent different categories and their corresponding values. Bar charts are particularly useful for comparing values among different categories and identifying the highest or lowest values. They provide a clear visual distinction between data points and allow for easy comparison. On the other hand, a line chart is ideal for representing trends and changes over time. It connects data points with lines, enabling users to observe patterns, fluctuations, and relationships between variables. Line charts are commonly used

to analyze time series data, such as tracking the progression of Covid-19 cases over weeks or months. The comparison of calendar heatmap, bar chart, and line chart is shown in Figure 14.

While all three visualization methods serve different purposes, the calendar heatmap stands out due to its ability to provide an intuitive and comprehensive overview of data by utilizing colors to represent intensity or magnitude. It condenses information and allows users to quickly identify patterns and variations in the data.

The calendar heatmap is a powerful visualization technique that effectively presents data in a grid-like format. It utilizes colors to represent the magnitude, intensity, or density of a particular variable across different categories or time periods. The key advantage of a calendar heatmap is its ability to provide a quick and holistic understanding of data patterns. By using a color spectrum, it allows users to easily identify areas of high or low values, patterns, and outliers. This is especially valuable when working with large datasets or when exploring complex relationships between multiple variables.

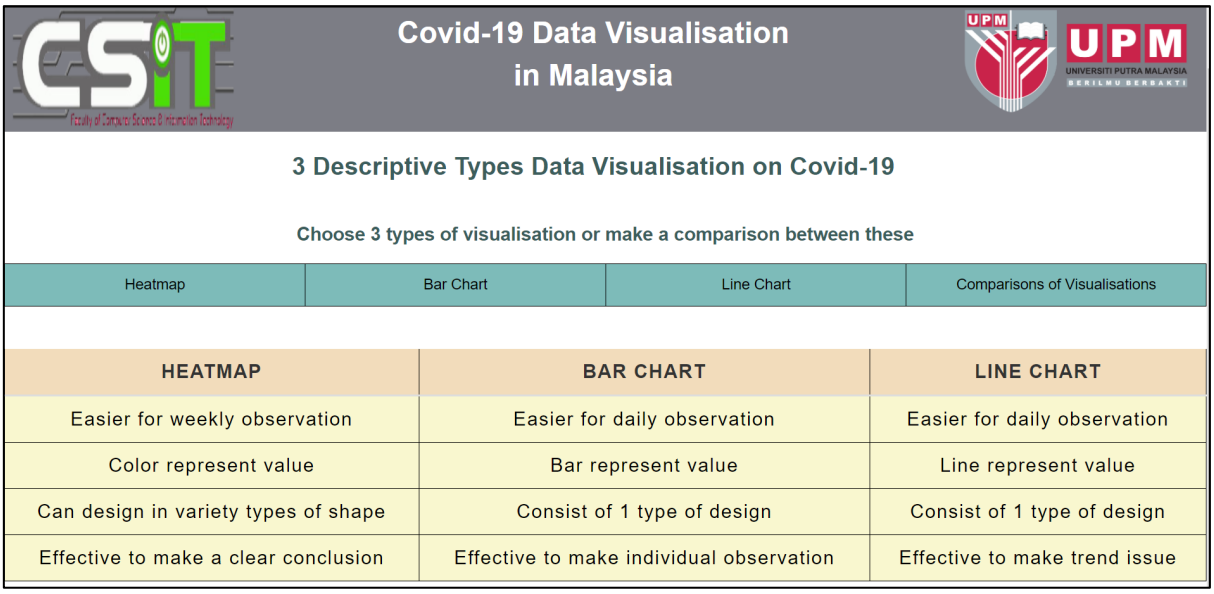

Fig. 14. Result on Display (Comparisons of Visualizations)

4. Conclusions

A Covid-19 descriptive visualization approach that uses calendar heatmaps is proposed to visualize the spread of the virus over time and space. The approach was designed to be a comprehensive web-based platform that offers multiple data visualizations for tracking Covid-19 cases in Malaysia.

One of the significant achievements of the proposed approach was the creation of a descriptive analytics tool that effectively communicates the daily Covid-19 cases and associated death rates in Malaysia. By offering calendar heatmap visualization, the current state of the epidemic is successfully conveyed and able to provide valuable insights.

However, it is important to note that the proposed approach has limitations. One notable weakness is its inability to capture and display the important dates and factors influencing the trends of Covid-19 cases, such as the impact of significant events or interventions. For instance, the approach fails to represent the effectiveness of measures like the Movement Control Order implemented in Malaysia on March 18th, 2020, which significantly contributed to the reduction in cases.

For future work, the proposed approach can be improved by incorporating important dates and contextual factors that influence the Covid-19 case trends. This would provide a more comprehensive understanding of the underlying dynamics and facilitate better decision-making for stakeholders involved in epidemic management. The proposed approach can also be used to visualize other types of pandemic, such as the spread of influenza or HIV/AIDS. The suggested improvements are targeted at providing a more comprehensive understanding of the underlying dynamics and facilitating better decision-making for stakeholders involved in epidemic management.

References

[1] Rana, Rashmi, Ravi Kant, Rohit Singh Huirem, Deepika Bohra, and Nirmal Kumar Ganguly. "Omicron variant: Current insights and future directions." *Microbiological Research* (2022): 127204. https://doi.org/10.1016/j.micres.2022.127204.

[2] Cascella, Marco, Michael Rajnik, Abdul Aleem, Scott C. Dulebohn, and Raffaela Di Napoli. "Features, evaluation, and treatment of coronavirus (COVID-19)." (2020). https://pubmed.ncbi.nlm.nih.gov/32150360/.

[3] CodeBlue. 2023. "MOH to Review Malaysia's Covid Epidemic Mid-June, after WHO Ends PHEIC Declaration." CodeBlue. May 8, 2023. https://codeblue.galencentre.org/2023/05/08/moh-to-review-malaysias-covid-epidemic-mid-june-after-who-ends-pheic-declaration/ .

[4] Shah, Ain Umaira Md, Syafiqah Nur Azrie Safri, Rathedevi Thevadas, Nor Kamariah Noordin, Azmawani Abd Rahman, Zamberi Sekawi, Aini Ideris, and Mohamed Thariq Hameed Sultan. "COVID-19 outbreak in Malaysia: Actions taken by the Malaysian government." *International Journal of Infectious Diseases* 97 (2020): 108-116. https://doi.org/10.1016/j.ijid.2020.05.093 .

[5] Sheng, Jie, Joseph Amankwah-Amoah, Zaheer Khan, and Xiaojun Wang. "COVID-19 pandemic in the new era of big data analytics: Methodological innovations and future research directions." *British Journal of Management* 32, no. 4 (2021): 1164-1183. https://doi.org/10.1111/1467-8551.12441.

[6] Roser, Max, and Hannah Ritchie. 2020. "Coronavirus Disease (COVID-19)." Our World in Data 1 (1). https://ourworldindata.org/coronavirus.

[7] Fernandez, Nicolas F., Gregory W. Gundersen, Adeeb Rahman, Mark L. Grimes, Klarisa Rikova, Peter Hornbeck, and Avi Ma'ayan. "Clustergrammer, a web-based heatmap visualization and analysis tool for high-dimensional biological data." *Scientific data* 4, no. 1 (2017): 1-12. https://doi.org/10.1038/sdata.2017.151.

[8] Comba, Joao LD. "Data visualization for the understanding of COVID-19." *Computing in Science & Engineering* 22, no. 6 (2020): 81-86.https://doi.org/10.1109/mcse.2020.3019834.

[9] So, Mike KP, Agnes Tiwari, Amanda MY Chu, Jenny TY Tsang, and Jacky NL Chan. "Visualizing COVID-19 pandemic risk through network connectedness." *International Journal of Infectious Diseases* 96 (2020): 558-561. https://doi.org/10.1016/j.ijid.2020.05.011.

[10] Khanam, Fahima, Itisha Nowrin, and M. Mondal. "Data visualization and analyzation of COVID-19." *Journal of Scientific Research and Reports* (2020): 42-52. https://doi.org/10.9734/jsrr/2020/v26i330234.

[11] Our World in Data. 2022. "COVID-19 Data Explorer." Our World in Data. 2022. https://ourworldindata.org/explorers/coronavirus-data-explorer.

[12] WorldOMeter. 2023. "Coronavirus Toll Update: Cases & Deaths by Country." Worldometer. 2023. https://www.worldometers.info/coronavirus/.

[13] The New York Times. 2023. "Track Covid-19 in the U.S.," March 22, 2023, sec. U.S. https://www.nytimes.com/interactive/2023/us/Covid-cases.html.

[14] Dong, Ensheng, Jeremy Ratcliff, Tamara D. Goyea, Aaron Katz, Ryan Lau, Timothy K. Ng, Beatrice Garcia et al. "The Johns Hopkins University Center for Systems Science and Engineering COVID-19 Dashboard: data collection process, challenges faced, and lessons learned." *The lancet infectious diseases* 22, no. 12 (2022): e370-e376. https://doi.org/10.1016/S1473-3099(22)00434-0.

[15] Ismail, Izzatul Syahirah, Siti Hajar Aishah Samsudin, Muhammad Adam Sani Mohd Sofian, and Hamidah Jantan. "COVID-19 Vaccination Data Visualization: Issues and Challenges." In *2022 International Visualization, Informatics and Technology Conference (IVIT)*, pp. 301-308. IEEE, 2022. https://doi.org/10.1109/ivit55443.2022.10033417.

[16] Hamzah, F. Binti, C. Lau, Hafeez Nazri, Dominic Vincent Ligot, Guanhua Lee, Cheng Liang Tan, M. K. B. M. Shaib, Ummi Hasanah Binti Zaidon, Adina Binti Abdullah, and Ming Hong Chung. "CoronaTracker: worldwide COVID-19 outbreak data analysis and prediction." *Bull World Health Organ* 1, no. 32 (2020): 1-32. https://doi.org/10.2471/blt.20.255695.

[17] Tao, Grace, Yuan Miao, and Sebastian Ng. "COVID-19 topic modeling and visualization." In *2020 24th International Conference Information Visualisation (IV)*, pp. 734-739. IEEE, 2020. https://doi.org/10.1109/IV51561.2020.00129.

[18] Wong, Man Kai, Donald J. Brooks, Juniorcaius Ikejezie, Marta Gacic-Dobo, Laure Dumolard, Yoann Nedelec, Claudia Steulet et al. "COVID-19 mortality and progress toward vaccinating older adults—World Health Organization, worldwide, 2020–2022." *Morbidity and Mortality Weekly Report* 72, no. 5 (2023): 113.https://doi.org/10.15585/mmwr.mm7205a1.

[19] Santermans, Eva, Emmanuel Robesyn, Tapiwa Ganyani, Bertrand Sudre, Christel Faes, Chantal Quinten, Wim Van Bortel et al. "Spatiotemporal evolution of Ebola virus disease at sub-national level during the 2014 West Africa epidemic: model scrutiny and data meagreness." *PloS one* 11, no. 1 (2016): e0147172. https://doi.org/10.1371/journal.pone.0147172.

[20] Cifuentes, Myriam Patricia, Clara Mercedes Suarez, Ricardo Cifuentes, Noel Malod-Dognin, Sam Windels, Jose Fernando Valderrama, Paul D. Juarez et al. "Big Data to Knowledge Analytics Reveals the Zika Virus Epidemic as Only One of Multiple Factors Contributing to a Year-Over-Year 28-Fold Increase in Microcephaly Incidence." *International Journal of Environmental Research and Public Health* 19, no. 15 (2022): 9051. https://doi.org/10.3390/ijerph19159051.

[21] Awotunde, Joseph Bamidele, Rasheed Gbenga Jimoh, Idowu Dauda Oladipo, Muyideen Abdulraheem, Taibat Bolanle Jimoh, and Gbemisola Janet Ajamu. "Big data and data analytics for an enhanced COVID-19 epidemic management." In *Artificial Intelligence for COVID-19*, pp. 11-29. Cham: Springer International Publishing, 2021. https://doi.org/10.1007/978-3-030-69744-0_2

[22] Costa, Carlos J., and Joao Tiago Aparicio. "A Methodology to Boost Data Science in the Context of COVID-19." In *Advances in Parallel & Distributed Processing, and Applications: Proceedings from PDPTA'20, CSC'20, MSV'20, and GCC'20*, pp. 65-75. Springer International Publishing, 2021. https://doi.org/10.1007/978-3-030-69984-0_7.

[23] Our World in Data. 2022. "COVID-19 Data Explorer." Our World in Data. 2022. https://ourworldindata.org/explorers/coronavirus-data-explorer.

[24] Meyer, Miriah, Tamara Munzner, Angela DePace, and Hanspeter Pfister. "MulteeSum: a tool for comparative spatial and temporal gene expression data." *IEEE transactions on visualization and computer graphics* 16, no. 6 (2010): 908-917. https://doi.org/10.1109/tvcg.2010.137.

[25] Urribarri, Dana K., and Martín L. Larrea. "A visualization technique to assist in the comparison of large meteorological datasets." *Computers & Graphics* 104 (2022): 1-10. https://doi.org/10.1016/j.cag.2022.02.011.

[26] Kwon, Sejeong, Meeyoung Cha, Kyomin Jung, Wei Chen, and Yajun Wang. "Prominent features of rumor propagation in online social media." In *2013 IEEE 13th international conference on data mining*, pp. 1103-1108. IEEE, 2013. https://doi.org/10.1109/icdm.2013.61.

[27] Khade, Anindita A. "Performing customer behavior analysis using big data analytics." *Procedia computer science* 79 (2016): 986-992. https://doi.org/10.1016/j.procs.2016.03.125.

[28] Bostock, Mike. 2013. "D3.Js - Data-Driven Documents." D3js.org. 2013. https://d3js.org/.

[29] D3. 2023. "Calendar." Observable. May 31, 2023. https://observablehq.com/@d3/calendar/2?intent=fork.

[30] BBC. 2020. "Covid Map: Coronavirus Cases, Deaths, Vaccinations by Country." BBC News, January 29, 2020. https://www.bbc.com/news/world-51235105.