VIET NAM NATIONAL UNIVERSITY HO CHI MINH CITY
INTERNATIONAL UNIVERSITY
INFORMATION TECHNOLOGY - DEPARTMENT OF COMPUTER SCIENCE

Data Science and Data Visualization

Project report
INTERACTIVE VISUALIZATIONS OF TRACKING COVID-19 PANDEMIC SITUATION

Lecturer: Dr. Mai Hoang Bao An
Team: H3K-Solution

No	Student name	Student ID	Contribution
1	Nguyen Thi Mai Huong	ITITIU19128	25%
2	Vu Ba Hung	ITITIU19126	25%
3	Truong Duy Khang	ITITIU19137	25%
4	Ho Thi Thu Hoa	ITITIU19120	25%

Ho Chi Minh City, May 20th, 2023

TABLE OF CONTENTS

LIST OF FIGURES

1 Introduction

1.1 Background and Motivation

Covid-19 is a global pandemic caused by the new coronavirus. This epidemic has affected many countries around the world, causing hundreds of millions of infections (767,364,883 confirmed cases) and millions of deaths (6,938,353 deaths)[1]. The COVID-19 pandemic has affected people's lives and economies in many countries worldwide, with economic crises taking place, a series of businesses and factories closing, and people panicking. With many variants born, COVID-19 became a concern in many countries around the world.

Since the beginning of the COVID-19 pandemic, it has become a global emergency when the number of infections and deaths has increased rapidly and is difficult to control. Many countries cooperate in research to find the optimal solution to the above problem. In that situation, data science and analytics play an important role in making appropriate health policies to prevent the transmission of the pandemic. Data visualization makes it easier and more accurate to track the number of infections, deaths, and recoveries, from which many optimal solutions have been proposed. Up to now, the world has controlled the pandemic, but it has not completely disappeared and there is still a risk of a resurgence with new variants. So our team built interactive visualizations of tracking the COVID-19 pandemic situation project to observe and analyze the pandemic.

1.2 Problem statement

Although, the COVID-19 pandemic has been controlled and the world is recovering from the consequences caused by the disease in a positive way. However, the World Health Organization (WHO) released important information on March 16 on the situation of variants of *SARS-CoV-2*, WHO listed many variants as multiple variants of concern (VOCs) such as the Omicron viruses. And a number of variants with worrying symptoms have been identified by WHO since the beginning of the COVID-19 pandemic. So we cannot neglect the prevention and monitoring of the COVID-19 situation, ***our project was built with the purpose of*** creating a reliable website that collects data related to the epidemic, and after cleaning, visualize these. Finally, provide many visualizations of too much data to customers, it is convenient to research and monitor the situation and trends of the disease.

1.3 Objectives and scopes

The ***main objective*** of the interactive visualizations of tracking the COVID-19 pandemic situation project is to visualize data using images such as graphs, charts, and maps to present data about the COVID-19 pandemic such as the number of infections, number of deaths, number of recoveries, population by each country in the world. The purpose of the project is to build an application program to reduce the manual work for observing and accessing pandemic data.

We collected data on the COVID pandemic and processed it to produce a highly reliable web page, and then presented that data in an easy-to-understand manner using graphs, charts, and maps. Our ***system scope*** aims for a visualization website that helps customers have an overview of the pandemic situation by capturing and making decisions that are appropriate to the situation. Customers will easily observe, compare, and evaluate data from many countries around the world with charts.

2 Literature review

2.1 Data visualization advantages in tracking situations

In the age of big data, data visualization becomes popular with obvious advantages[2] such as:

- Intuitive: People are attracted to colors and patterns in clothes, art and culture, architecture, etc. The data is also the same - colors and images make it easy for them to see and understand the data.
- Simple data sharing: Information is shared in an accessible, understandable way for many different audiences.
- Better Analysis: Its easier to spot the relationships and patterns within a data set when the information is presented in a graph or chart. Thus, people are more readily able to understand and draw conclusions from these.
- Quicker Decisions: It's easy for people to observe, compare, and evaluate data with graphs, charts, and maps. From there make decisions quickly and in accordance with each situation.

–> In the context of the outbreak of the COVID-19 pandemic, data visualization makes a great contribution to making health policies to prevent the pandemic such as social distancing, the "5k" method, or monitoring the effectiveness of the COVID-19 vaccine.

2.2 COVID-19's Data Visualizations

The data on the COVID-19 pandemic is collected, cleaned, encrypted, and used by the World Health Organization (WHO) using data visualization to represent:

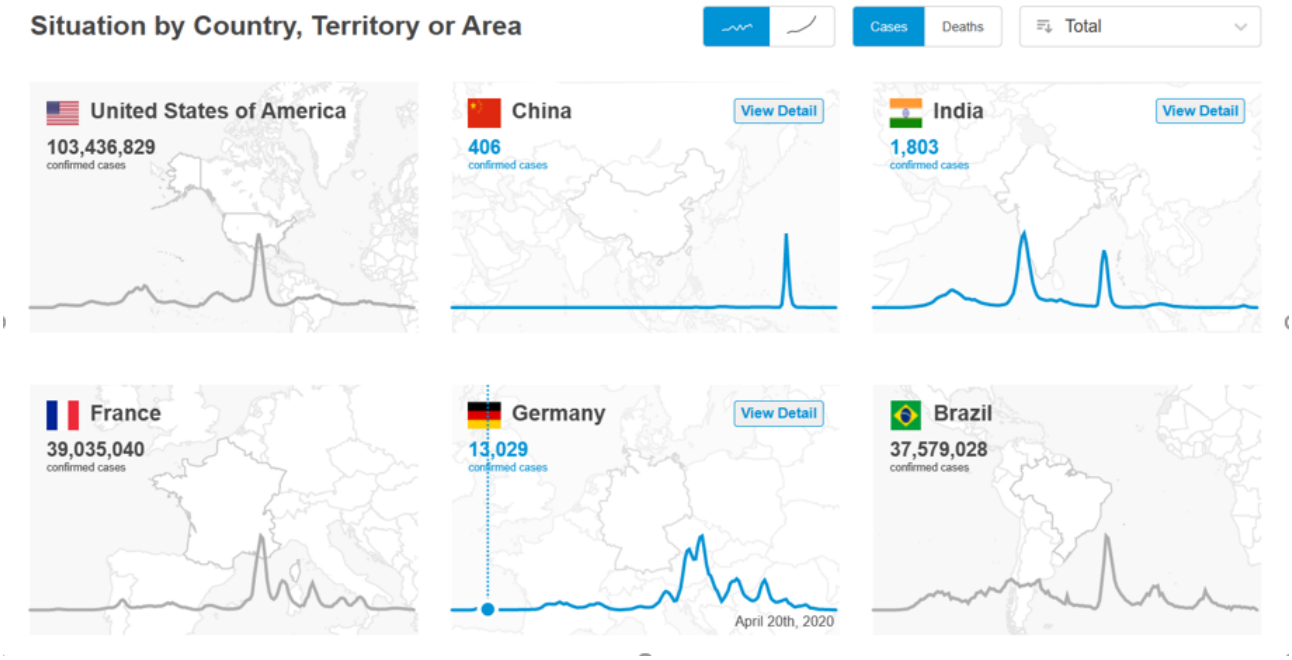

Figure 1: The line chart represents COVID-19 confirmed cases of each country

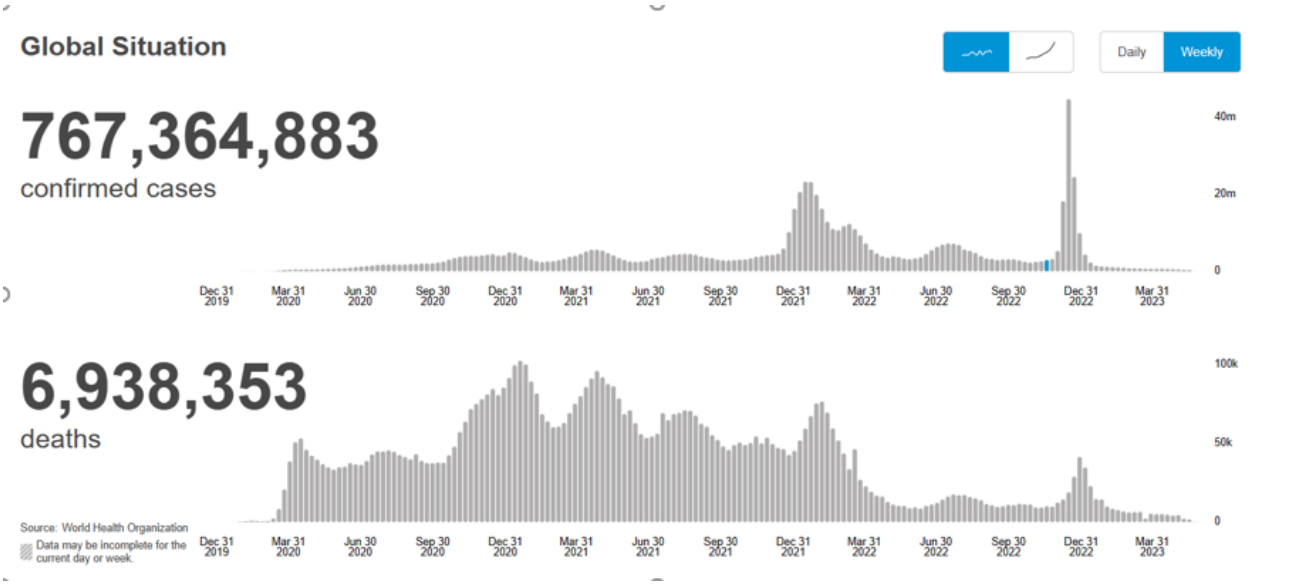

Figure 2: The line chart represents COVID-19 worldwide confirmed cases and a total of deaths on a daily/weekly basis

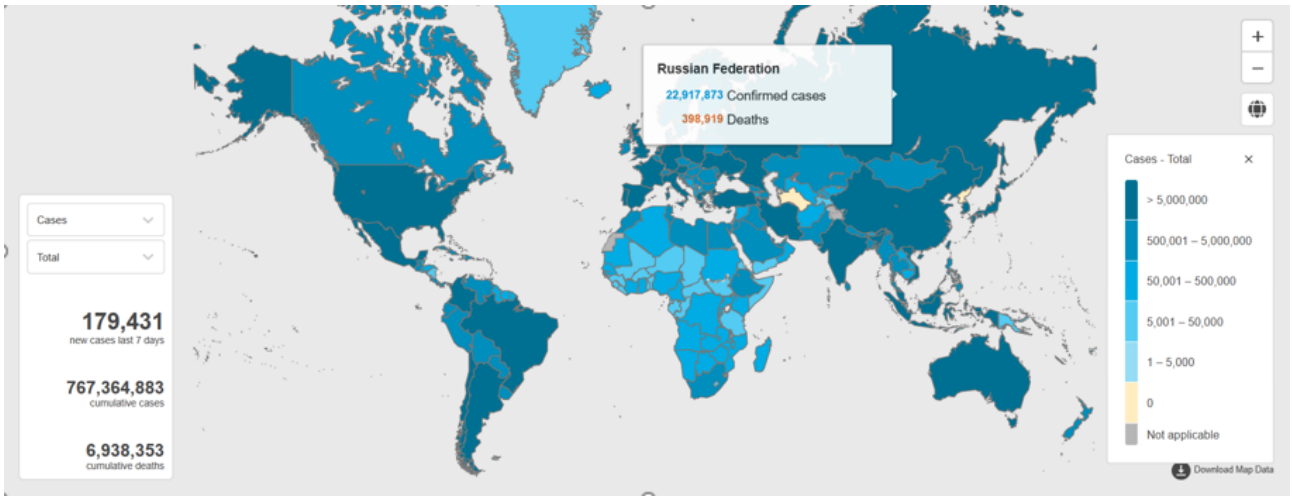

Figure 3: The geographical chart that focus colours by total cases

–> The World Health Organization (WHO) is a global and highly influential organization with high credibility. WHO uses visualizes data in order to present data through many images and this has made it easy for people and governments to track and access information quickly and accurately. From there, people have an overview and understand the importance of COVID-19 prevention as well as the benefits of vaccines. As a result, the disease is quickly controlled and new variants are easily detected to provide optimal measures.

2.3 Similar COVID-19 Analyses

The COVID-19 pandemic is a global problem that directly affects the economic and social crises of many countries around the world, so many research groups and organizations have created direct systems to visualize data to present data relative to the pandemic, for research purposes, and discuss optimal solutions for the problem above.

The World Health Organization (WHO) is one of the leading organizations in applying data visualization for the COVID-19 pandemic. In order to people can understand pandemic data meaning easily through images graphs, charts, and maps.

3 Project structure and dataset

Based on our main scope, we had to get covid data from Open Disease Data API [1] to shape the needed dataset that will be mapped into our geographical map. Whenever the user hovers the mouse over any country region, the map will pop up additional information related to that country. If he/she clicks on that country, the Bar Chart and

Line Chart will automatically show its country information in the chart container. The user can also choose the way data is displayed or add more countries to the chart by selecting a country in a bar chart. He/ she also can delete all/ some countries that he/ she doesn't need to show in the bar chart. We will put our state diagram right below to let you easier imagine the way our project works.

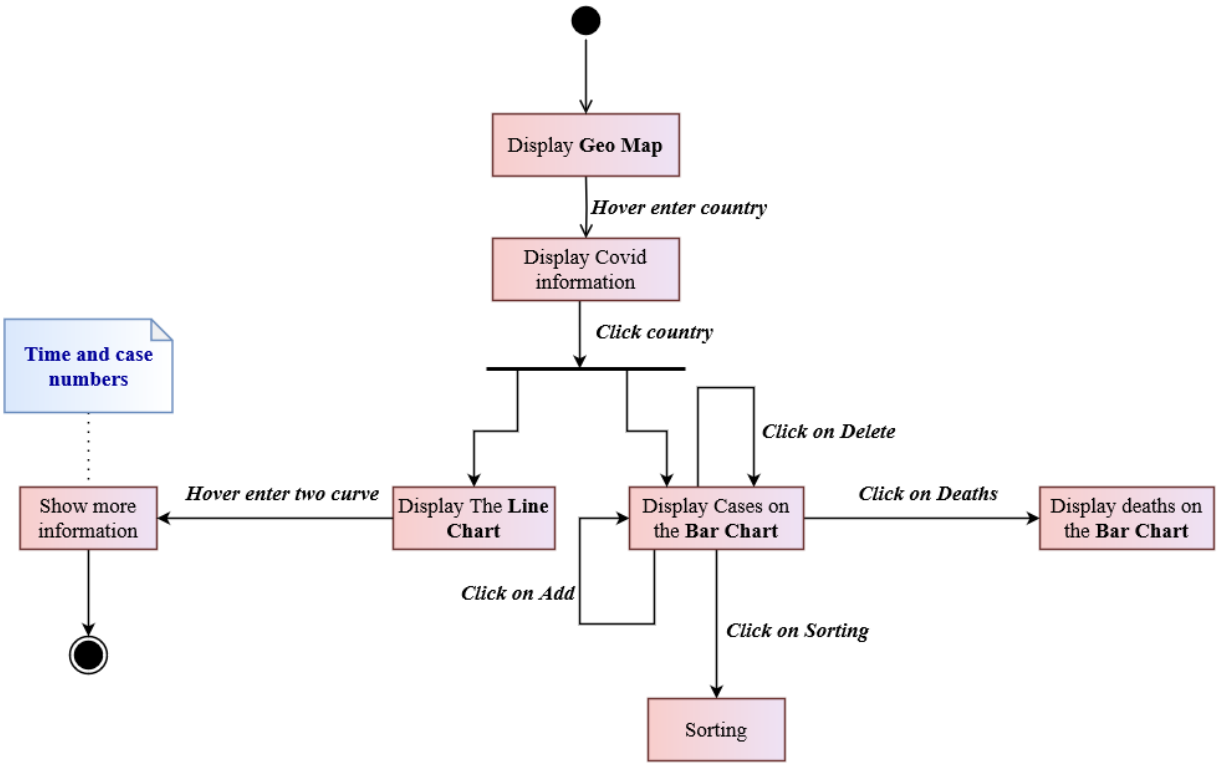

Figure 4: State diagram of our project's main flow process

Our project includes 3 charts:

- **Geographical chart**: Show covid information of each country in the world and average cases over 1 billion people.
- **Bar chart**: Compare the number of cases and deaths between selected countries.
- **Line chart**: Show the number of cases and deaths in a country within a day(the last day).

4 Demo prototypes

4.1 Overall project

At the beginning of initiation charts, it will show readers the overview geo map at the top and the bar chart with empty content in below:

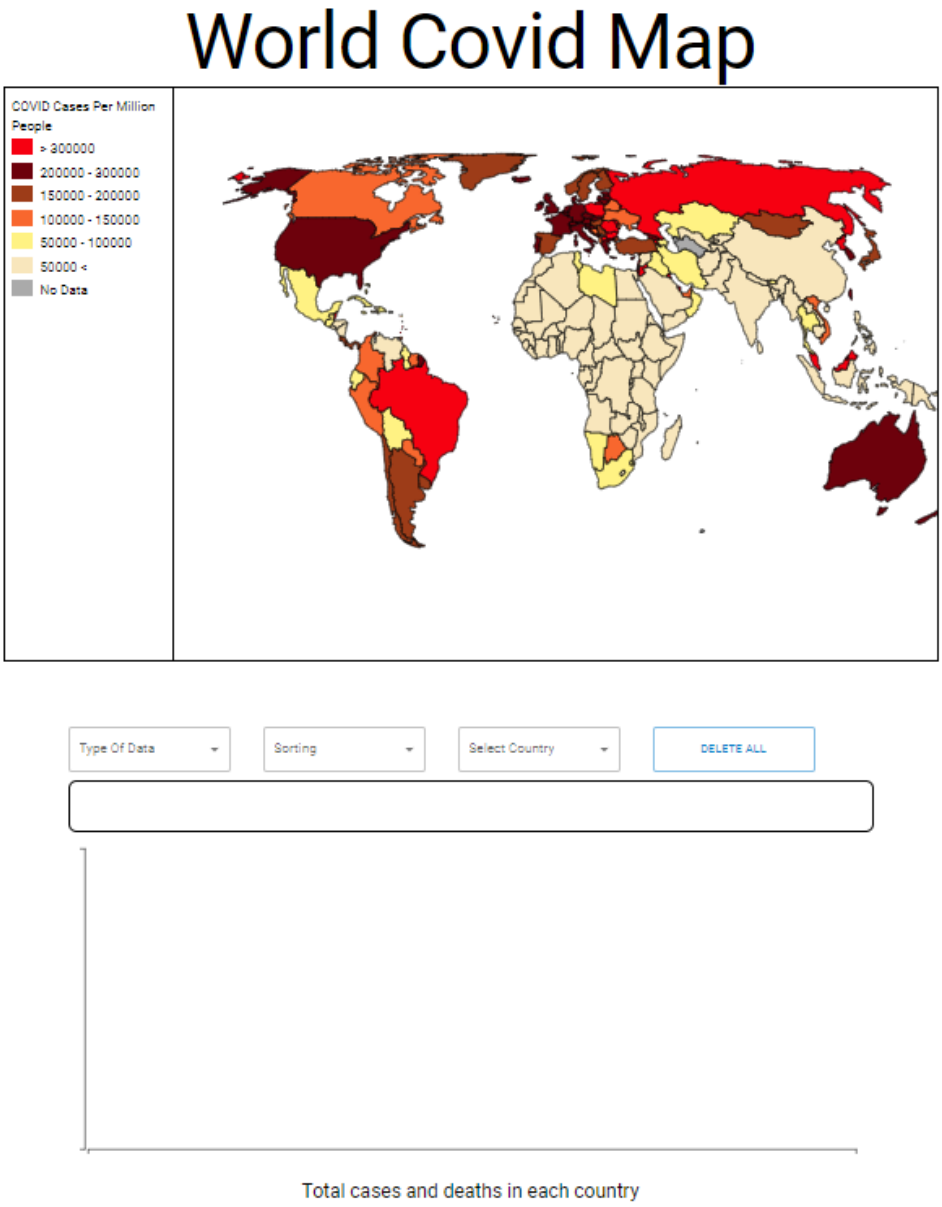

Figure 5: The overview of our charts

When users click on any region it will show the additional details of that country, and automatically be inserted into the bar chart and line chart to show total cases and deaths as well as the number of confirmed cases respectively:

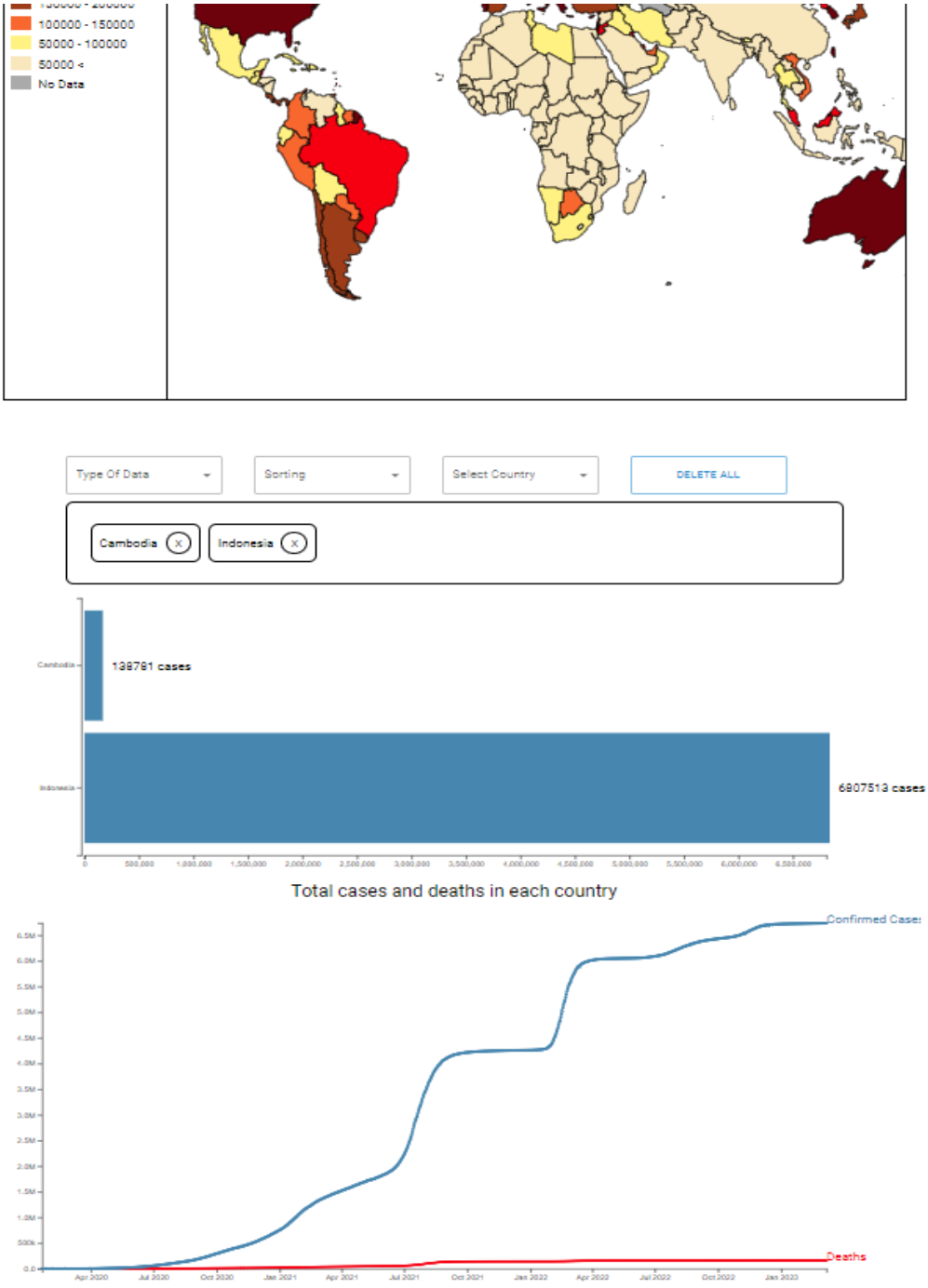

Figure 6: Charts are updated when the user interacts with the first geo chart to select a country(ies)

We will demonstrate more details of each chart in the next sections.

4.2 Visualization of the number of COVID Cases in the geo chart

In order to let readers quickly capture the overview of how serious the COVID-19 cases situation worldwide is, we have implemented the covid world map that colored regions by the case numbers.

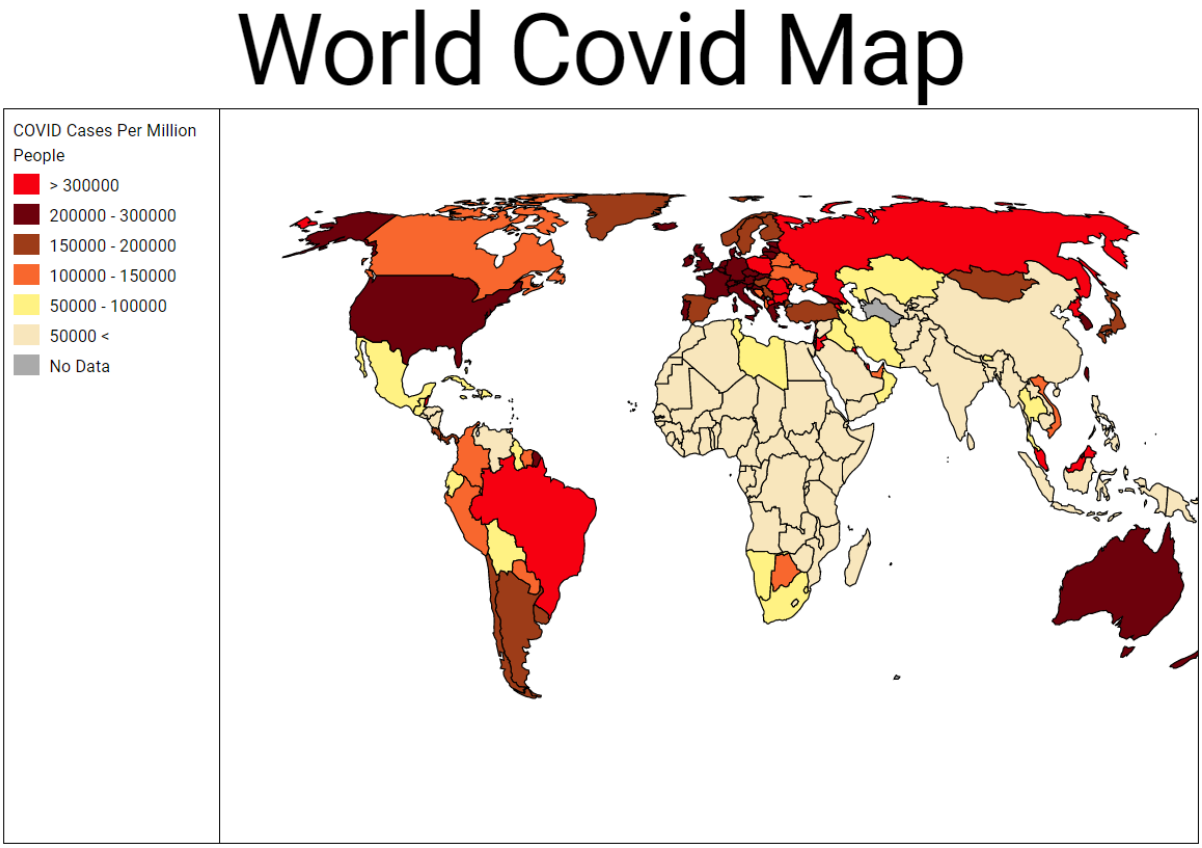

Figure 7: The overall covid world map

Our map can be zoomed in, zoomed out, swiped around, hovered to a region to see more details of that country.

Figure 8: Displaying additional information of USA related to covid situation when the mouse hovers to this country

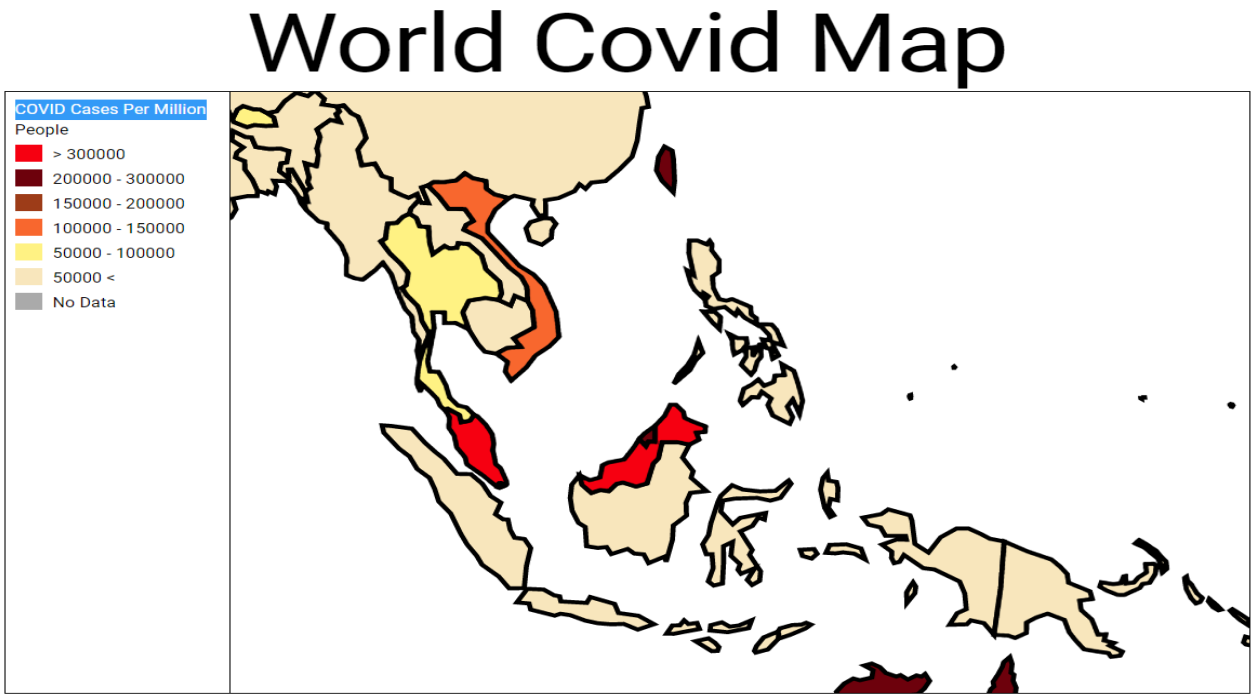

Figure 9: The map when mouse zooming it out to clearly see several small countries

4.3 Covid-19 visualization via the bar chart

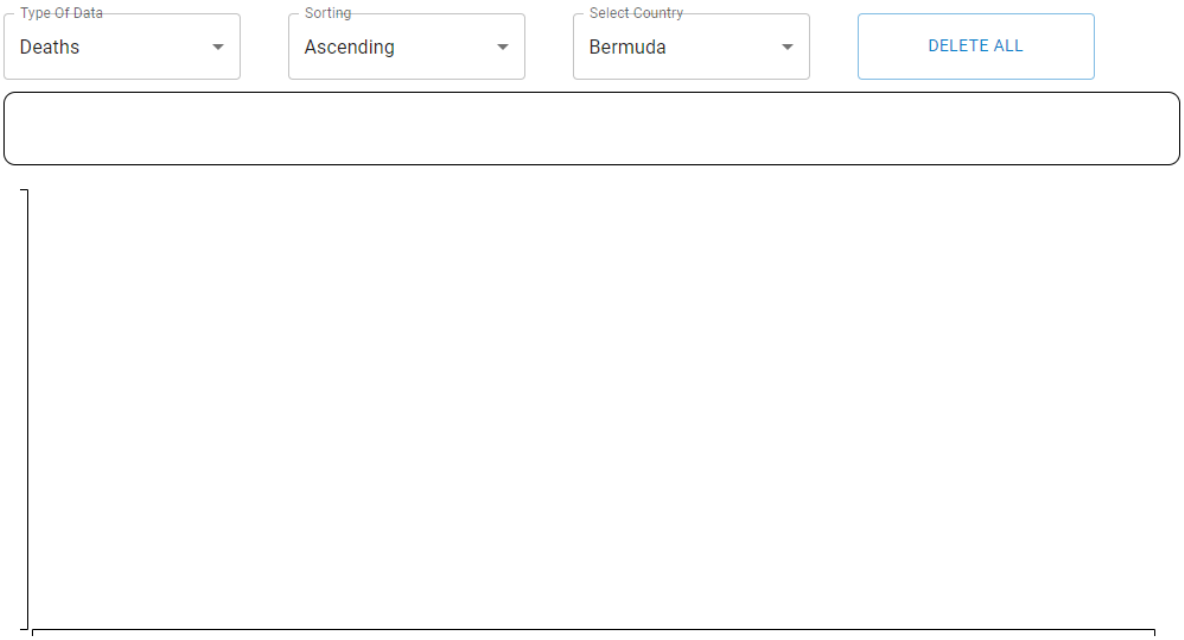

Figure 10: The empty bar chart with no selected country

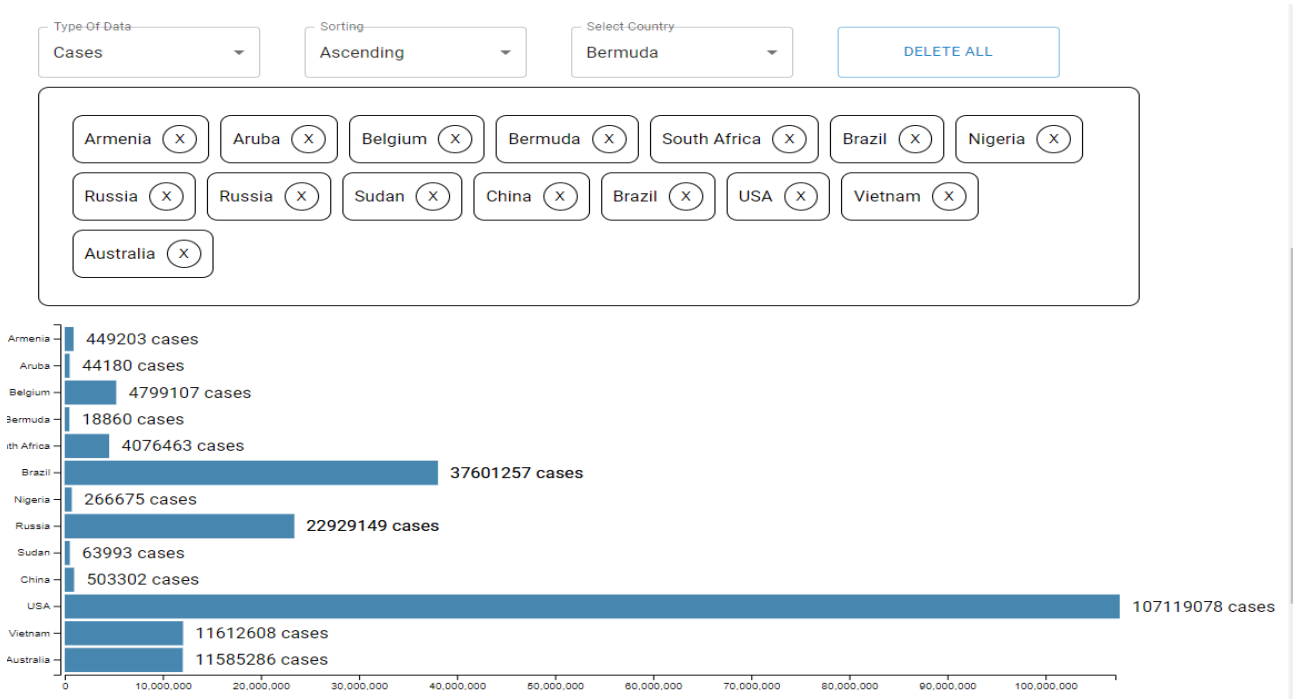

Figure 11: The bar chart display cases of selected countries by added period order

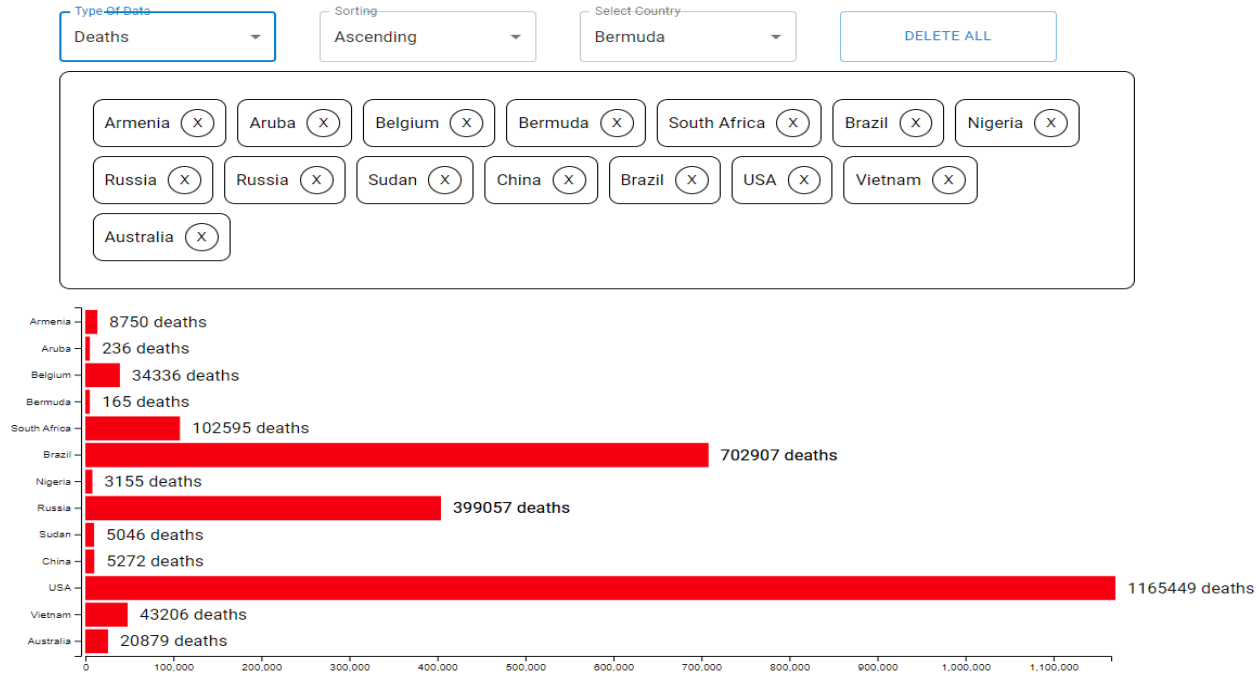

Figure 12: The bar chart displays death of those above countries when the user chooses Deadth as the type of data

Figure 13: The bar chart displays case number of those above countries sorted in decreasing order

The bar chart also can sort data in increasing order as well but we will show it later in the presentation demo part.

4.4 The line chart

As we already mentioned in the previous parts, it is used to show the total cases and deaths in a country.

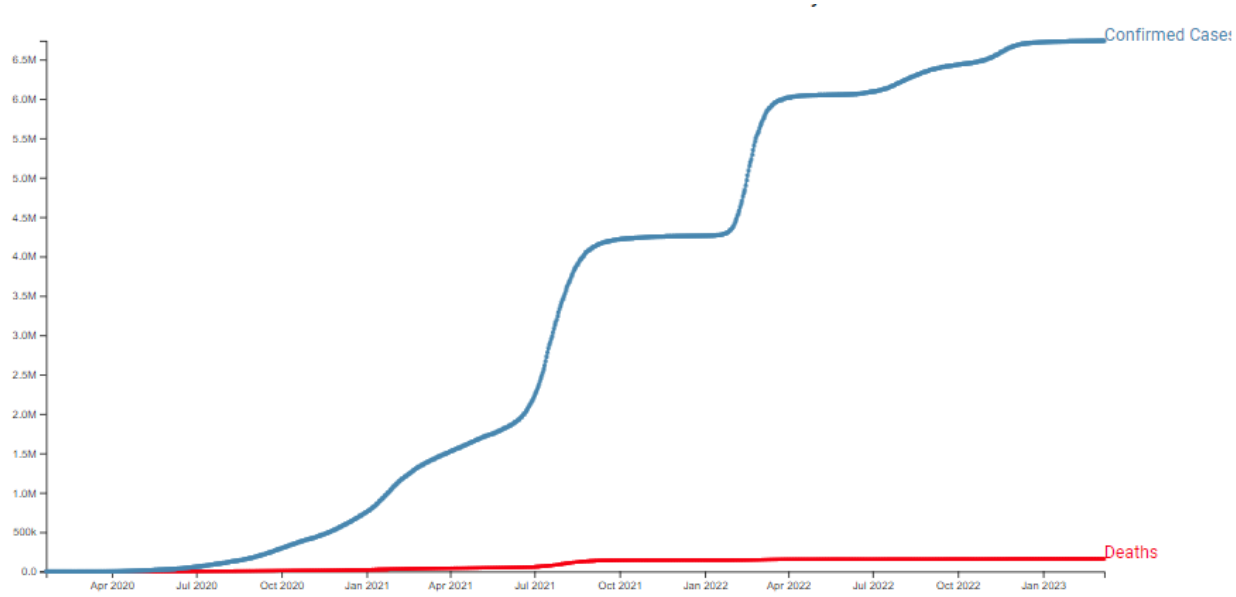

Figure 14: The line chart displays total cases and deaths in Viet Nam in given period

5 Conclusion and Discussion

5.1 List of accomplished work

Our project has primarily reached the scope defined in the first section when we initiate our project, that aims to let users quickly get the worldwide covid-19 epidemic situation, just mainly focuses on the cases and deaths via the geographic map, and can get top severe counties that affected by covid or make a comparison of several regions via sort function in our bar chart.

Accomplishing the project in a team not only gives us the chance to collaborate with teammates to enhance our communications skills and expand our individual network with fellow students but also is a great opportunity to help us exchange knowledge related to data visualization, which leads to the better understanding learned lessons in the theory class.

The project also is the best chance to help us practice with several libraries such as react-simple-maps, which can significantly contribute to our coding experience enhancement that serves for our future job as software engineers.

Moreover, via the project we learn the techniques to examine data, work with data like sort, filter, etc, and visualize them as much efficient as possible.

5.2 Strength and Weakness

5.2.1 Strength

Our team has gained a lot of knowledge and experiments during the project. All team members are consensus with a common working direction through each stage and get in mind the big picture of the project in the very first meetings. Due to the differences in coding ability, we don't clearly separate tasks to assign for each member but work together step-by-step in order to help others in improving coding skills and to be sure of understanding the achieved result. Each of us always strives to complete the common and help others to better understand what lessons they have learned and how to apply them to the project.

In the first few weeks of making prototype designs, we determined the development method, the topic, and its objective as well as the scope, and defined a list of tasks that need to be done in time. We also hold meetings on a weekly basis to track task progress and discussed more if a member has any trouble or obstacle. This makes us much easier to grasp the group's work and complete our tasks.

5.2.2 Weakness

Although we had a quite good project management methodology and team collaboration, applying theoretical knowledge to code-based project visualization is quite difficult for us. It required a lot of effort to continuously revise lab sections related to line charts, bar charts, etc, and also practice the chart libraries offered by react framework in order to get the hang of it and build the workable demo.

We also realize that building a tracking COVID-19 pandemic situation that gets good user experience requires a lot more than just what we have done! We need to dig into integrating several aspects related to covid-19 cases, building real-time platforms, implementing more interactive charts, and so on. Furthermore, our methods of parsing and mapping data to the charts are also needed to optimize to make them less complicated but still work smoothly.

5.3 Conclusion

To sum up, our project has been somewhat successful in terms of showing the world-wide COVID-19 pandemic situation. We tried to build our project in a simplified direction that just serves enough required functions that are targeted in the project objective and scope.

If there is an opportunity to develop our project in the near future, we will research several efficient methods for mapping the dataset into the application, making it real-time. Besides, we also implement more functions for the project and build both web-based and mobile apps for it.

6 References

[1] ***WHO Coronavirus (COVID-19) Dashboard***, WHO (COVID-19) Homepage: https://covid19.who.int/

[2] ***What are the advantages and disadvantages of data visualization?***, tableau: https://www.tableau.com/data-insights/data-visualization/advantages-di

[3] ***disease.sh - Open Disease Data API***: https://disease.sh/