CS411 Project Report

SQL Database Design - Daily Life with COVID-19

Jiahe Li, Qianyan Shen, Luting Lei

2022 FALL

Daily Life with COVID-19

1. Please list out changes in directions of your project if the final project is different from your original proposal (based on your stage 1 proposal submission).

We intended to use past data from each country to predict future trends in our initial design. Users can select regions and times to see how the model predicts outbreak development, as well as the model's accuracy based on past data. To find a suitable model, we will combine some of the most recent research findings.

We are now concerned about whether we or our close friends are at risk of contracting the virus. Therefore, we designed registration for users to focus on specific country or region.

2. Discuss what you think your application achieved or failed to achieve regarding its usefulness.

The most difficult aspect of designing the database and functionality was the prediction component. We intended to use past data from each country or region to predict future trends in our initial design. However, when we began designing, we discovered that the data for these predictions had already been implemented in many databases and were not particularly innovative for us. Predictions based solely on past data are also not the most accurate. As a result, we needed to consider more specific data. Certain cities, for example, are important transportation hubs with a high volume of people coming and going. Therefore, our actual design abandons this aspect of prediction.

3. Discuss if you changed the schema or source of the data for your application

Finding relevant data for each of our boards is relatively simple. The WHO database can be used to obtain basic information such as the number of confirmed diagnoses. Many well-known journals, including the Lancet, Nature, and Science, have separate sections for COVID-19 news and research results, making it easier for us to collect data. We can also learn about social media debates on Twitter, Facebook, and other platforms. We choose social media from various countries to make our work more comprehensive.

In actual design, our database is overburdened by the large amount of data. As a result, we decided to keep only the more relevant data, such as the most recent data.

4. Discuss what you change to your ER diagram and/or your table implementations. What are some differences between the original design and the final design? Why? What do you think is a more suitable design?

We did not fully consider what type of data the user wanted to query when designing the ER diagram, so the number of tables was very small. We included a lot of content in the final design, such as the user's "area of interest."

5. Discuss what functionalities you added or removed. Why?

In our initial design, we wanted to leverage historical data from each country or region to forecast future trends. When we started designing, we discovered that the data for these predictions had already been implemented in many databases and were not particularly novel for us. Predictions based solely on historical data are also not the most precise. As a result, we had to look at more specific data.

6. Explain how you think your advanced database programs complement your application.

SQL-based procedures and triggers are two of our advanced features. Triggers are used to categorize risk levels. Simply put, high-risk areas are those with a

high number of diagnoses or deaths. We use a procedure to implement the risk situation of the registered users' focused areas and present it to them.

7. Each team member should describe one technical challenge that the team encountered. This should be sufficiently detailed such that another future team could use this as helpful advice if they were to start a similar project or where to maintain your project.

Jiahe Li: Designing ER charts and thinking about how these data actually relate to each other in real life is complex. In our practical attempts, we have found that there are many relationships that cannot be represented simply. At the end of the semester, we learned about Neo4j, which inspired us that if we could use NoSQL for optimization, this would be a great option.

Qianyan Shen: We had never learned about front and back ends before, and designing as a group led us to have some problems in interfacing the different modules. For example, how to apply the code of SQL query practically on the web page.

Luting Lei: When designing database operations, we need to pay extra attention to user permissions. For example, users cannot make changes to our core data. We should give the user the appropriate operation rights according to the reality.

8. Are there other things that changed comparing the final application with the original proposal?

A similar question has been answered earlier. We have mainly modified the prediction section.

9. Describe future work that you think, other than the interface, that the application can improve on

In class, we learned about Neo4j, which inspired us the architecture of NoSQL database. The intimate contact between people is a significant factor in controlling the covid-19 outbreak. As a result, if we can record which areas

people visited at when times, we may estimate the probability of catching the virus. Of course, if we include this, we must make certain that this data is not exposed to all users for privacy concerns.

10.Describe the final division of labor and how well you managed teamwork.

We ended up with a completely different group split than we had considered at the beginning. First of all, we had one less group member, so our three members had a heavier task. In addition, we were more inclined to work together than to work individually on our own tasks.