

CPT103 Assessment 3:

The Pandemic Outbreak Monitoring System Database

Updates of Coursework Requirements

Please remember to check the updates and common Q&A on the LearningMall about this coursework. This document might be updated one or two times if certain specifications are not clear enough. All updates will appear here or be **highlighted** in this document. Remember to check emails about any updates.

Database Specifications

The Lukewarm Kingdom is a famous tourist place for people who want to enjoy sunny beaches and lovely weather. The recent outbreak of pandemics caused a lot of panic in this country. To better prevent and control outbreaks, the government decided to develop a pandemic monitoring and management system (PMMS). You have been appointed to design a database based on the requirements of the PMMS and provide detailed documentation about its design and usage. You will need to apply all the knowledge throughout the course. Only database design is involved, no C/Java programming is needed for this project.

The PMMS is designed to be used across the whole Lukewarm Kingdom. The system is split into two functional components. The first component records viral test results within the whole Lukewarm Kingdom. The second component is designed to record and manage citizen travel history. Citizen travel records are used to predict potential areas of virus spread if someone infected has been wandering around these areas. The following two sections will describe the background information and requirements of these two components in detail.

Viral Test Recording

Viral tests are done in hospitals and all test reports will be managed centrally by the PMMS. Currently, there are 12 hospitals in the Lukewarm Kingdom capable of doing viral tests. But more hospitals will be added in the future.

All Viral test reports follow the same template shown in Figure 1. A viral test report starts with the basic information about the patient who took the test, which is then followed by the sample type and test result. The example only shows one sample type, but your design must support future sample types. The footnote “Coughid-21 is a newly...” appeared at the end of the report is the description for the sample type “Coughid-21”. All samples of the same type will have the same description.

Central Lukewarm Kingdom Hospital			
Name	Jianjun Chen	Sex	Male
Age	70		
Mobile	8008101818	Sample Type	Coughid-21
Sample Result			
Positive			
Sample Collect Time	2020/10/1 13:27	Sample Test Time	2020/10/1 14:50
Doctor:	Jun Qi	Report Time	2020/10/1 17:20
* Coughid-21 is a newly identified type of virus this year, all patients tested to be postive should rest well and avoid going outside			

Figure 1 Example Viral Test Report

The temporal information of viral tests is also recorded in the report. The process of viral tests starts with the sample collection then followed by the sample testing, where the positivity or negativity of the sample is identified. Once the results are confirmed, the report will then be generated with the doctor's name. To make things simple, we assume that citizens who are tested to be positive will be immediately sent to stay in the hospital until he becomes healthy again. Any viral tests that occurred during the treatment will not be recorded in the PMMS.

All information listed in the example above must be included in your database design. The knowledge of normalisation is needed. For the report of this part, you need to list all functional dependencies of the elements in the sample report and normalise it into 3NF. You can add more attributes, but a detailed explanation must be given. Note that only patients who did the viral test should be recorded in the database. This database is not designed to store the information of all citizens in the country.

Travel History Tracking

The Lukewarm Kingdom is split into 5 regions: north, south, east, west and central regions. Each region is further divided into sub-areas called "districts". The government of the Lukewarm Kingdom has set three risk levels for districts: high, mid and low. If a district has one or more positive cases staying for more than 24 hours, it will be considered a high-risk district. Districts that have positive cases within 1 week (even count people just passing by) are considered to be mid-risk. Districts with no positive cases within 1 week are low-risk areas. Districts are the smallest "area of interest" when monitoring pandemic outbreaks. For example, if positive cases were found in a certain district, the whole district will be isolated from other places.

The key infrastructure used in travel history tracking is mobile base stations (BS). Each base station stores its GPS information that describes its location within the Lukewarm Kingdom. When a mobile phone with a SIM card enters the signal range of a BS, it will automatically register itself to this BS using the ID (assume it is the phone number) of its SIM card. A BS is also aware when a mobile phone moves out of its signal range.



Figure 2 Mobile base station

A special program installed in BSs will automatically upload tracking data into your database. The program has not been fully developed yet because the underlying database still needs to be designed by you. But this special program will be allowed to use one or more INSERT or UPDATE statements to upload tracking data to your database.

To protect citizens' privacy, the database for travel history monitoring should not include too much of citizens' personal information. According to the laws of the Lukewarm Kingdom, only the following 4 fields can be obtained from these BSs: GPS location, SIM card ID, SIM card connection time (when entering the signal range of BS), SIM card disconnection time (when leaving the signal range of BS). To keep things simple, this coursework also makes the following assumptions:

1. The signal coverage areas of BSs are mutually exclusive (no overlapping area of BSs signal)
2. All areas inside the Lukewarm kingdom are fully covered by BSs (no blind spot).
3. All mobile phones are always on.
4. One BS only belongs to one district. a BS will never service two or more districts at the same time.
5. Phone numbers are unique for each citizen. One citizen can only have one phone number.
6. **Everyone has a mobile phone.**

The format of GPS location must be searched and learned by yourself. You will be asked to explain your GPS format in your report.

Important Use Cases

This section lists some very important use cases of the PMMS. Your database design is expected to satisfy all of these use cases. As a result, if you have designed your database based on the previous two sections but later realised that it cannot satisfy all use cases below, you need to revise your design until they can be done. *Keep in mind that all use cases below can be satisfied with a single SELECT statement (unless specified otherwise, sub-queries in a query is not counted as another query).* Test data is crucial for this part. Data must cover items that matches the goal and items that

do not match the goal. The relationships between data items must be explained clearly in your report.

Use case 1: A person can potentially get infected if he was in the same district with someone. The government requires that, if someone is tested to be positive, all people in the same district as him in the past 48 hours (before the positive report is published) need to take viral tests. Assume that a person called Mark was tested to be positive at 19:30 on 09-Oct-2021. **Mark's telephone number is 233636.** Please write a query that can get the phone numbers of all citizens who will potentially get infected because of him.

Use case 2: Please first clearly describe the format of GPS locations. The format must be a valid format that is used in real life. Then mimic what happens to your database when a user moves into the range of a base station and then moves out one hour later by listing all SQL statements involved in the process.

Use case 3: The Lukewarm Kingdom wants to find out the hospitals that can do viral tests efficiently. The report generation time is calculated using report time minus the sample test time. Please write a query to find out which hospital has the least average report generation time.

Use case 4: List the phone numbers of all citizens who did two viral tests with the time window from 2021-10-03 00:00 to 2021-10-05 00:00. The two viral tests must have a gap time of at least 24 hours (at least 24 hours apart).

Use case 5: **List the high-risk, mid-risk and low-risk districts using one query. High-risk districts should be listed first, followed by mid-risk districts and then low-risk districts.** Example:

district_name	risk_level
Centre Lukewarm Hillside	high
Lenny town	high
Glow Sand district	mid
Raspberry town	low
Bunny Tail district	low

Use case 6: List all positive cases found in the district called "Centre Lukewarm Hillside" on 2021-10-04. The result should include the names and phone numbers of people tested to be positive.

Use case 7: Calculate the increase in new positive cases in the district called "Centre Lukewarm Hillside" on 2021-10-05 compared to 2021-10-04. The result should show a single number indicating the increment. If there are fewer new positive cases than yesterday, this number should be negative.

Use case 8: Assume that the spread rate of a virus is calculated by dividing the total number of people that were in the same district as the positive case with 48 hours (calculated in **use case 1**) by the total number of people among them that later confirmed to be infected within 14 days. Again,

assume that a person called Mark (telephone number is 233636) was tested to be positive at 19:30 on 09-Oct-2021 and he is the only person in the country that had coughid-19 at that time. Please write a query that calculates the spread rate of the virus.

Extended Use Cases

In addition to the use cases in the previous section, you are required to design 10 more use cases and implement them. Details are explained in the report template.

Your Tasks and Marking Criteria

You are required to write a detailed report explaining your database design for the PMMS. For the database design, you need to make an ER diagram that fits **one single page**. Please do not split the diagram into several pages.

Please strictly follow the report template and answer all of the questions in it. The ER diagram only constitutes a small part of your marks. The majority of your marks come from the quality of your discussions and ideas.

You also need to submit the SQL script containing all CREATE TABLE statements that create the tables. **Your SQL script and your ER diagram must match (attributes = columns, entities = tables, 1:m relationships = foreign keys).** **The script must also contain the data and SELECT statements of use cases.**

The final design must be in the third normal form (3NF) and has no M:M relationships. You can make some assumptions about the data if they are not specified in the requirements. For example, you can assume address occupies up to 200 characters, or people's names can be up to 30 characters long. You need to indicate these decisions in the report. The SQL script should also include the SQL statements for use cases.

Item	Marks
Database design and explanation (report quality)	60% (out of 100%)
Important Use Cases	24%
Extended Use Cases	16%

Coursework Submission

The submission deadline is 23:59, 13th of December and the cut-off date is 23:59, 17th of December. Each day of late submission leads to a 5% deduction of the total marks of this coursework. This assessment contributes 70% towards the overall module grade.

The format of CW Submission on the LearningMall:

Everything should be packed in a zip file. The filename of the zip file should follow this format:
`YourID_givenName_familyName.zip`

For example: `123456_Jianjun_Chen.zip`

In the zip file:

1. A document called "report.pdf", which is your reference document. Please do not submit word documents directly. Export the word document to PDF first (Word, WPS and Libreoffice all support this). 1 mark will be deducted if the file format is not PDF.
2. A script called "script.sql", which contains all of your SQL statements. A template is provided.