Final Exam - BTM 382

Thank you for your application for the role of Database Administrator (DBA) at OurCompany Inc.! We are currently reviewing applicants and would like to congratulate you on making it past the first round of screening. As part of our company's hiring policies, we require that applicants for this role submit an assessment of their competencies. The assessment will comprise of the following deliverables, to be produced and returned to OurCompany within 48 hours of issuance;

1 During your role as DBA you will be required to perform requirements analysis to determine what type of information needs to be stored, and understand the operational needs of the database. For this assessment we have provided you with a transcript of an interview with a client, found in Appendix A. You will need to read this transcript and synthesize the needs of the database based on the client's responses to the interview questions. Based on your understanding of the client's needs, produce a complete ERD including entities, relationships with cardinality & optionality, PKs & FKs, etc. This deliverable must be produced within MySQL Workbench. Submissions for this deliverable:

      - A PDF of your Complete ERD - Using MySQL Workbench, when you have your EER Diagram open, go to File>Export>Export as Single Page PDF.

2 Using the sample data provided in the attached excel document, go through the steps of normalization, showing your work through each step. Clearly identify any and all keys, partial dependencies, transitive dependencies and full functional dependencies amongst the attributes in the database. The sample data is from the client's previous database, so it's contents do not necessarily represent what will need to be stored for the new database.

      - The deliverable for this assessment is a document containing your work for normalizing the database based on the sample data. The document can be PDF, word, excel, etc as long as it is legible.

3 Based on the ERD you produced in Deliverable 1, implement this database in MySQL. Provide the DDL script required to implement this database. In addition to creating the schema, create 10 rows of made-up data for each table in your database.

      - The deliverable for this assessment is a .sql file containing the DDL script for your ERD from 1. Save it as ########_CREATE.sql, but replace ######## with your student ID.

      - Include another .sql file containing the DML statements required to populate your tables with data. Save it as ########_INSERT.sql, but replace ######## with your student ID.

4 Your manager has issued a request for several reports to be developed. This work order is listed in Appendix B. You will need to codify these requirements into SQL queries.

- Provide the code for each query along with a screenshot of the results.

5 Assuming acceptance of your database design and implementation, the client has some follow-up questions found in Appendix C.

- Read through the questions and provide a written answer to each. The total word count for all questions should not exceed 500 words.

Appendix A

The following transcript is of an interview with the head of operations at BioMed Pharmaceuticals ("The Client"). This interview was conducted as part of the feasibility analysis and requirements assessment phase of the contract. BioMed Pharmaceuticals is a manufacturer of prescription medications, and they've contacted you ("The Contractor") to assist in building an updated database for their operations.

1. Greetings, so I understand you'd like to update the database for your pharmaceutical operations. Can you tell me a bit more about your business?

"Certainly. So we produce prescription medications for a huge variety of ailments and distribute these nation-wide. We purchase our raw materials from a network of suppliers, after which we process the raw materials to transform them into the final product, which is then sold to a network of distributors. "

2. Interesting, can you explain a bit more about this supply-chain? Who are the suppliers and what sort of information are you trying to keep track of for these?

"So the suppliers are just other large companies who produce raw materials. A given supplier will often produce multiple different kinds of raw materials, and a given raw materials can be produced my many different suppliers - for instance, one of our anti-stress & alertness drugs, Focusyn, frequently used by students during exam period, is produced using a blend of the same raw material; some is produced locally here in Canada but other parts of the blend are imported. As another example, another drug product we sell to foreign governments, NK-Ultra, is produced using a raw material which is only provided by one supplier locally. Generally we'd like to know the company name, address, name of contact person, contact phone number, Industry Rating for each supplier. Ratings can be from 1 to 5, indicating an unreliable supplier or a great supplier. Raw materials should have their name stored in the database and the country of origin."

3. Great, and what about the distributors? How does it work exactly? Do they interact with the suppliers? What information do you want to store about the distributors?

"No, the distributors do not interact with the suppliers directly. The distributors purchase products from us at wholesale prices and then sell them to retail outlets. Some distributors specialize with one single product, while others will diversify and sell a multitude of our products. For distributors, well we'd like to know a lot of the same information as for suppliers, but also include the number of years that the distributor has been working with us. We often like to send out thank-you letters and gifts for long time customers, so it's important we know which distributors have been buying from us for the longest."

4. Nice. You mentioned retail outlets, do you want to keep track of them as well, and if so, what do you want to store?

"Glad you asked. As part of our analytics we do want to collect data on retailers to understand more thoroughly how our product is being received by the end-user. Again for retailers we want to store a lot of the same information as for suppliers and distributors. For retailers specifically,

in addition to the general company information, we'd like to record the Retail License Number. The retail license numbers change every few years after a renewal. During potential litigations or recalls, ultimately we are responsible for verifying the authenticity of the retailers, so we must do our own due diligence and verify the Retail License Numbers. If an unlicensed retailer is caught selling our products, our brand can be negatively affected, so we need to actively verify that our retailers are up to code. Retailers can sell our products and may have purchased them from a single distributor or multiple distributors."

5. What about your products? What information do you store about these?

"Of course, the products! So our pharmaceuticals are made by processing raw materials, and this is done by one of our Processing Facilities. We want to know the product name, price, and the illness that it is used to treat."

6. Okay interesting, so you have your own processing facilities? What sort of information do you want to store for them?

"Well, in addition to the general company information we store for the other members of the supply chain, we also want to know the Facility License Number. These license numbers change every few years when the application is renewed and the facility passes inspection. In events of recalls, governments will want our license numbers promptly to verify that we are up to code."

7. Ok great. Is there anything else you want to keep track of in this database?

"Yes, a lot actually. Given the nature of our products and the health industry, we feel we need to have extreme oversight over the transfer of our products, from supplier to customer. Moreover, our products can't simply be bought from a retailer. Retailers need to see a prescription before providing a customer with any of our products. Customers need to get prescriptions from a Doctor. For both Customers and Doctors, we want to know the names, phone numbers and addresses. For customers, we want to know the gender, blood type, age, nationality, etc. Typical demographic information so we can run analyses based on which customers consume which types of drugs. For doctors, we need to record the Institution from which they got their Medical Degree. A prescription will have a date, quantity, and the doctors notes (ie. take 2 before bedtime)."

8. Impressive, I bet you'll be able to perform some complex analyses with all this information. I think this is plenty to work with, unless there's anything else you'd like to add?

"No, I agree this should be good to begin with. Let's see what you come up with. Thanks!"

Appendix B

In order for your database implementation to be accepted by the client, it needs to meet a series of acceptance criteria. These include a variety of reports that the database must be able to perform. You will be required to codify these reports into SQL. The reports demanded by the client are listed below;

1. In the event of a product recall, we want to know all the contact names for everyone involved. For a specific product name, list all the retailers' contact names who sold it, names of customers who have taken it, doctors' names who prescribed it, the facility contact name where it was processed, the names of raw materials included in it, and the contact names of the suppliers who provided those materials.

2. Perform an RFM analysis to determine which customers have filled prescriptions more recently than average, which customers have filled prescriptions more frequently than average, and which customers have spent more on their most expensive prescription than average.

   Based on this information, we might be able to further analyze trends in prescriptions with customer demographics. For customers who scored 3/3 for their RFM analysis, show their demographic information in a view.

3. Create 3 additional queries of your choosing. (Try to impress the client with the complexity of your queries.)

Appendix C

Assuming acceptance of your final implementation of the client's database, the client has some follow-up questions. Based on your answer you may be eligible for future contracts. Your total answer to both questions should not exceed 500 words.

1 "Considering the size and breadth of data we intend to collect and store, we're concerned that NoSQL may be a better choice for our database than MySQL. Can you justify why MySQL is a a better choice for our implementation than NoSQL? In which situation is each most appropriate to be used? If not using NoSQL, how could we address the issue of big data within our database?"

2. "Suppose we wanted to modify our operational database schema, which is optimized for transactions, to better support decision making and business intelligence, what changes would you recommend to the current schema?"