

Database Management 604 Final Project

Project Summary:

As InsureYours, our company seeks to transform the Healthcare industry by providing patients with a sophisticated platform that assists individuals and healthcare providers in identifying the most suitable insurance coverage (Insurance Provider) based on the patient's gender, age and medical condition. As the number of health challenges continue to rise across the world, insurance prices mirror the same behavior. In order to keep individuals healthy when facing medical problems, we believe that patients should receive the most optimal insurance provider for their specific medical condition. By tailoring insurance recommendations to individual patient profiles, our aim is to ensure they can make informed decisions about the insurance coverage. Through our innovative approach, we believe that we could not only provide a solution for individual patients, but we could also revolutionize how patients are treated in hospitals around the world.

Database System:

Through our Healthcare Dataset, we will use our data to provide accurate insights and analysis for our patients using their gender, age and medical conditions. By using the data from Prasad Patil's dataset, we will generate insights and analysis to get a deeper dive on how we could make patients make the most optimal decision on their insurance provider. We created an Entity Related Diagram that expresses all of the relationships between values within certain tables. With our data we used the Extract, Transform, and Load (ETL) to pull the data from an online source to our database. Finally, we use a powerful visualization tool in Power BI to facilitate connectivity between the code and the insights we could generate through

visualizations. Since our data is being consistently updated through our ETL process, the incoming data will be reflected in our database, thus, limiting risk.

Data Processes:

The data that is being extracted would be stored in a SQL Server virtual database. Within our database, we created two main tables which are: Patient Information and Insurance Information. Within the Patient Information table we created columns that have specific data types which will store our data in its proper format. Those columns include PatientID(INT), Name(VARCHAR), Gender (VARCHAR), Age(INT), BloodType(VARCHAR), Medical_Condition(VARCHAR). For the Insurance information table, we have two columns which are InsuranceID and Insurance_Provider_Name(VARCHAR).

Hardware:

Having a certain storage system that can manage huge volumes of data while creating analysis is a big part of our business. For this, we believe that having a robust virtual machine would be imperative to the overall success of our company. Features such as 20 Cores, 80 GB of RAM, and 160 GB of Temporary Storage will cost \$5,942.46 per month and \$71,309.52 per year.

Data Entry/Analysis/Visualization:

Our database will provide a user-friendly interface for healthcare providers to enter patient details like gender, age, medical conditions. After the healthcare provider enters a patient's information, patients will receive a visualization consisting of transparent data providing the best insurance provider for their specific medical condition.

Requirements:

- Data used: We have taken the data from Prasad Patil's dataset. We will generate insights on how we would help patients make optimal choices in hospitals.

- Data Storage Requirements:

1. Data to be stored:

PATIENT INFORMATION:

- PatientID (INT)
- Name (VARCHAR)
- Gender (VARCHAR)
- Age (INT)
- BloodType (VARCHAR)
- Medical_Condition (VARCHAR)

INSURANCE INFORMATION:

- InsuranceID (INT)
- Insurance_Provider_Name (VARCHAR)

2. Data Entry/ Capture Process:

- Patient Registration: A user-friendly interface for healthcare providers to enter patient details like gender, age, medical conditions.
- Insurance Recommendation Input: After the patient enters their details, the best insurance provider will be recommended based on the patient's profiles.

3. Data Types and Validations:

- Data Types: We used appropriate data types for each field (like VARCHAR, INT, DATE) to ensure the consistency with the types of data collected.
- Validation Rules: We implemented validation rules to ensure data integrity. Validated age ranges, gender entries and insurance provider names.

4. Reporting Requirements:

Insurance Recommendations Report: Data of insurance recommendation and associated costs will be displayed.

References

Patil, P. (2023, October 31). *Healthcare dataset* 📊. Kaggle.

<https://www.kaggle.com/datasets/prasad22/healthcare-dataset>

Group Members: Lokeswar Kudumula, Reshma Attipatla, Swetha Manchukonda, Meghana
Puttiga, James Gilmore Jr,