Database Project

Topic- Liver Cancer Risk Assessment and the Monitoring

Vijay Bhaskar Reddy Vaka

Nikhil Kumar Gudapuri

Narendra Rayapati

Contents

[Topic Introduction 2](#_Toc102121526)

[Database Entity Relationship Diagram 3](#_Toc102121527)

[Database Relational Schema 3](#_Toc102121528)

[Primary and Foreign Keys 4](#_Toc102121529)

[Data Tuples in the Database 4](#_Toc102121530)

[SQL Queries Processing in the Database 5](#_Toc102121531)

[Database Infrastructural Diagram 8](#_Toc102121532)

[Databases and Differences and Selection of Database 8](#_Toc102121533)

[DBMS Queries 9](#_Toc102121534)

[Diagrams Documentation 9](#_Toc102121535)

[Database Model or ERD Model 9](#_Toc102121536)

[Network Diagram 10](#_Toc102121537)

[Database Infrastructural Diagram 11](#_Toc102121538)

[Data Types and schema 11](#_Toc102121539)

[Schemas 11](#_Toc102121540)

[Sequencing data 12](#_Toc102121541)

[Pseudo Code 12](#_Toc102121542)

[Data Tables 12](#_Toc102121543)

[Conclusion 14](#_Toc102121544)

[References 15](#_Toc102121545)

**Figures**:

[Fig 1: Database Entity-Relationship Diagram 3](#_Toc102121516)

[Fig 2: Database Relational Schema 3](#_Toc102121517)

[Fig 3: Database Infrastructural Diagram 8](#_Toc102121518)

[Fig 4: Database Queries 9](#_Toc102121519)

[Fig 5: Database or ERD 9](#_Toc102121520)

[Fig 6 Database Network 10](#_Toc102121521)

[Fig 7: Database Infrastructure 11](#_Toc102121522)

[Fig 8: Schemas database 11](#_Toc102121523)

[Fig 9: Sequencing data 12](#_Toc102121524)

[Fig 10: Pseudo Code for Query Processing 12](#_Toc102121525)

# Topic Introduction:

Cancer is the abnormal cell uncontrolled growth in the body that develops when the control mechanism in the body stops functioning in the person. So, in this case, abnormal cells start growing inside the body, and old cells in the body do not die and they start growing out of control which leads to the death of the person. Such cells might form a tissue mass that is called a tumor. The selected topic in the research for what the database system is required to develop is the Liver Cancer Risk Assessment and the Monitoring. The database system would help to maintain the data on liver cancer in the body and the risk probability of this cancer. This database development is aimed at monitoring cancer and managing treatments effectively.

# Database Entity Relationship Diagram

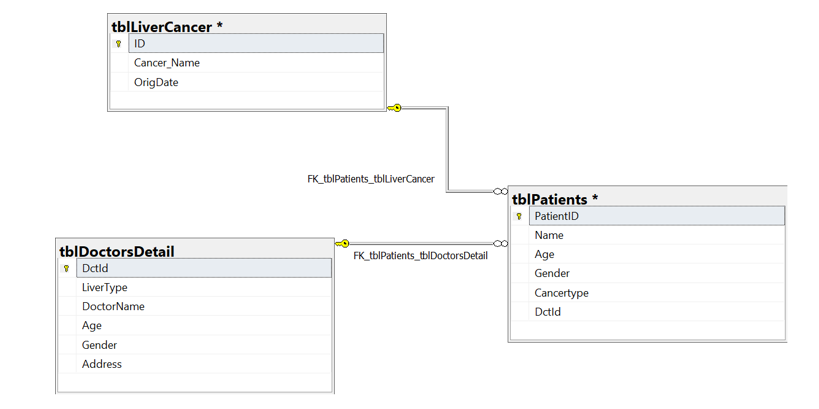
****

Fig : Database Entity-Relationship Diagram

While creating a database model, the very first requirement that comes in the process is the Entity Relation Model which helps to gather all kinds of data for the data management in the organization. There is a need to define the entities to insert into the system with the operations and attributes and the relationship as well in the entities (Chiang, et, al, 2019).

There are entities that are significant to record in the database such as patients, doctors’ detail, and liver cancer. The study is based on the risk of liver cancer and the effectiveness of Monitoring. When the database development would be planned then this would work effectively as all the entities are now collected and they would be processed with the attributes such as id, name, and others.

This helps the admin as well as the doctors to process any required data in the organization to manage and to see the real approach to the data management (Chiang, et, al, 2019).

# Database Relational Schema

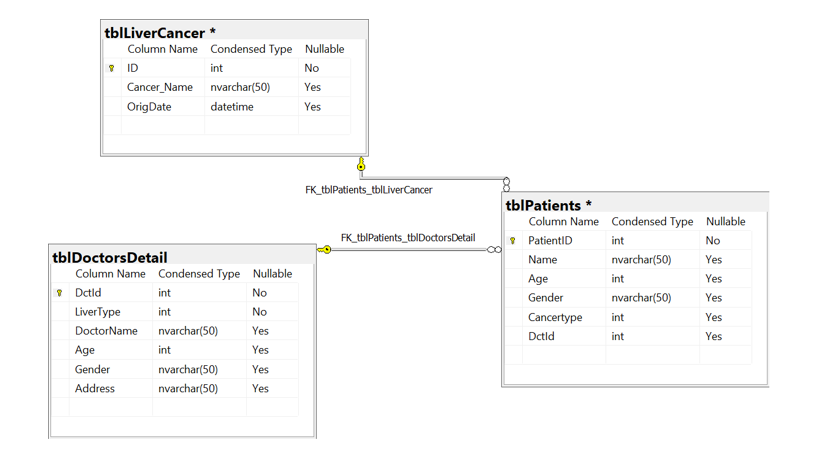


Fig : Database Relational Schema

When the data of the system is gathered or the attributes of the entities are collected then there is also a need to define the type of the data that gives the form of storing the data in the database. For example, if the Id is required to be captured in the database, then there is a need to define the data type such as integer form. This is called relational schema that create tables for the data that makes it easy to define all the data into the database in its real form whether integer or variable, dates or time. This is how the instructions are provided for the real outputs. This is required to perform and the above-shown diagram is presenting all the data with the tables that are containing its types effectively (Liang, et, al, 2019).

# Primary and Foreign Keys

**Patients’ data**

**Graphical user interface, application, table

Description automatically generated with medium confidence**

**Liver Cancer data**

**Table

Description automatically generated**

**Doctor Table**

**Table

Description automatically generated**

# Data Tuples in the Database

Table

Description automatically generated

# SQL Queries Processing in the Database

--Aggregate Functions

select max(age) as maximumAge from tblPatients

Rectangle

Description automatically generated with medium confidence

--Dates

select \* from tblLiverCancer where OrigDate > '1999-01-01'

Graphical user interface, text

Description automatically generated

--SQL Logical Operators like IN, AND, LIKE, NOT, OR, BETWEEN, EXISTS, etc.

select \* from tblLiverCancer where Cancer\_Name in ('Angiosarcoma','Hepatoblastoma')

Table

Description automatically generated

select \* from tblDoctorsDetail where Age > 30 and age < 70

Table

Description automatically generated

select \* from tblDoctorsDetail where DoctorName like '%a%'

Table

Description automatically generated

select \* from tblDoctorsDetail where DoctorName not in ('Robert')

Table

Description automatically generated

select \* from tblPatients where [Name] = 'John' or Age = 34

Table

Description automatically generated with medium confidence

select \* from tblPatients where Age between 0 and 50

Table, Excel

Description automatically generated

--Regular Expressions

SELECT Cancer\_Name FROM tblLiverCancer where Cancer\_Name like '[ka]%' ORDER BY Cancer\_Name;

Shape

Description automatically generated with low confidence

--Formula Expressions

select ID, Cancer\_Name,OrigDate from tblLiverCancer

Graphical user interface, text, application

Description automatically generated

--Subqueries

select \* from tblPatients where DctId = (select top 1 DctId from tblDoctorsDetail)

Graphical user interface

Description automatically generated with medium confidence

# Database Infrastructural Diagram

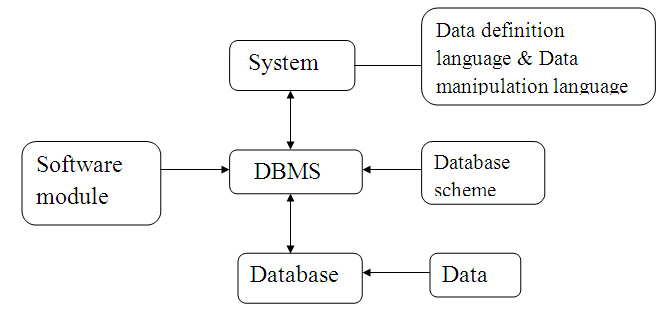


Fig : Database Infrastructural Diagram

This is an infrastructure model that is presenting the system with data definition, database scheme, data, database, DBMS, and the software modules.

While creating a database, all these are required effectively where the data in the database system can be recorded effectively. This is how the data is managed and results are achieved better as expected (Lv, et, al, 2022).

## Databases and Differences and Selection of Database

|  |  |  |  |
| --- | --- | --- | --- |
| No | Parameters | Relational Database (RDBMS) | Distributed Database (DBMS) |
| 1 | Data storage | The entire data is stored by the application but in the tabular form. These tables give primary keys. | All the applications are there to record the data in the file form. This gives navigation. |
| 2 | Normalization | There is the normalization of the data. | There is no normalization of the data. |
| 3 | Security | This provides integrity and this is more secure. | There are no security constraints for the data manipulations and data security. |
| 4 | Distributed database | In this, distributed systems are favored. | In this, distributed systems are not favored. |
| 5 | Table Relations | The entire data is stored by the application but in the tabular form. These tables give primary keys. Thus in the tables, there is a relationship. | All the applications are there to record the data in the file form. So, there are no relations in the tables. |
| 6 | User support | This database supports multiple users together at the same time. | This supports an individual user only. |

This is how the RDBMS is selected as this supports the multiple users and the data tables are received. RDBMS is found more secured comparatively and the results are effective data normalization helps to segregate the data where the security is high.

# DBMS Queries

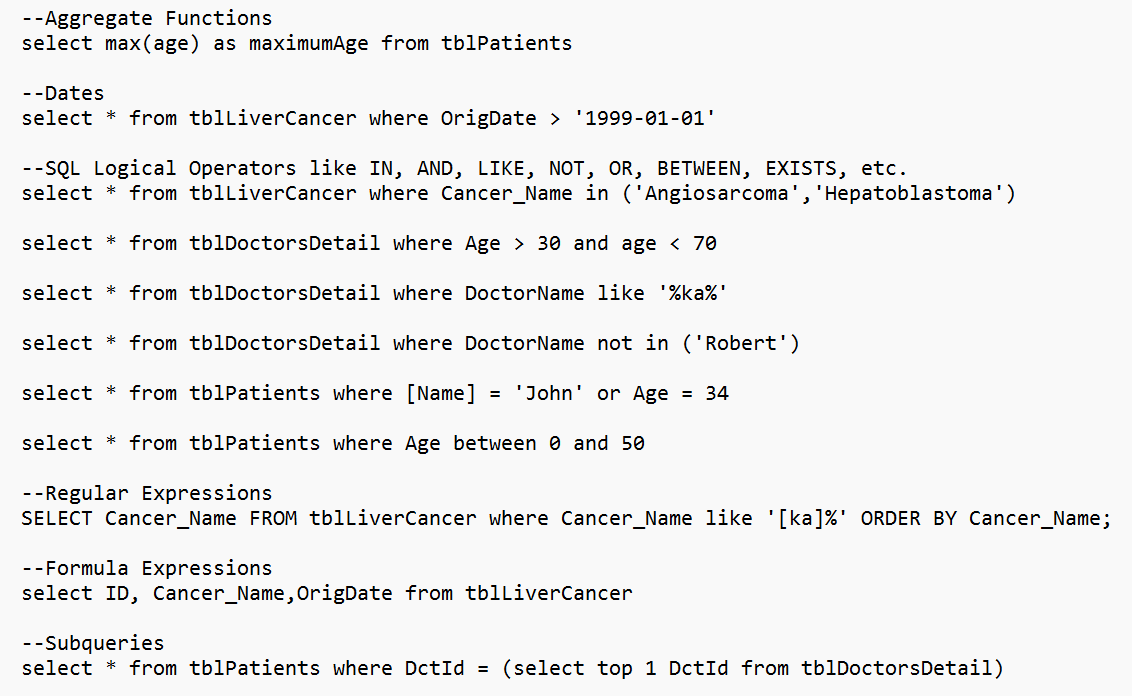


Fig : Database Queries

All these are queries that are processed in the database to extract data from the database.

# Diagrams Documentation

## Database Model or ERD

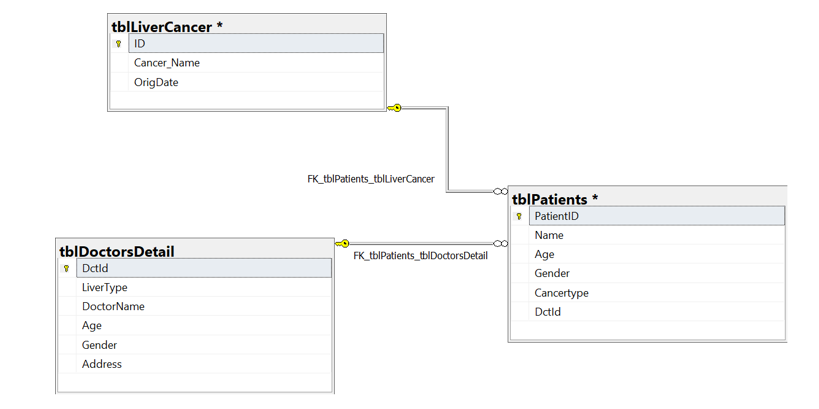
****

Fig : Database or ERD

There are entities that are significant to record in the database such as patients, doctors’ detail, and liver cancer. The study is based on the risk of liver cancer and the effectiveness of Monitoring (Cheung, et, al, 2019). When the database development would be planned then this would work effectively as all the entities are now collected and they would be processed with the attributes such as id, name, and others.

## Network Diagram

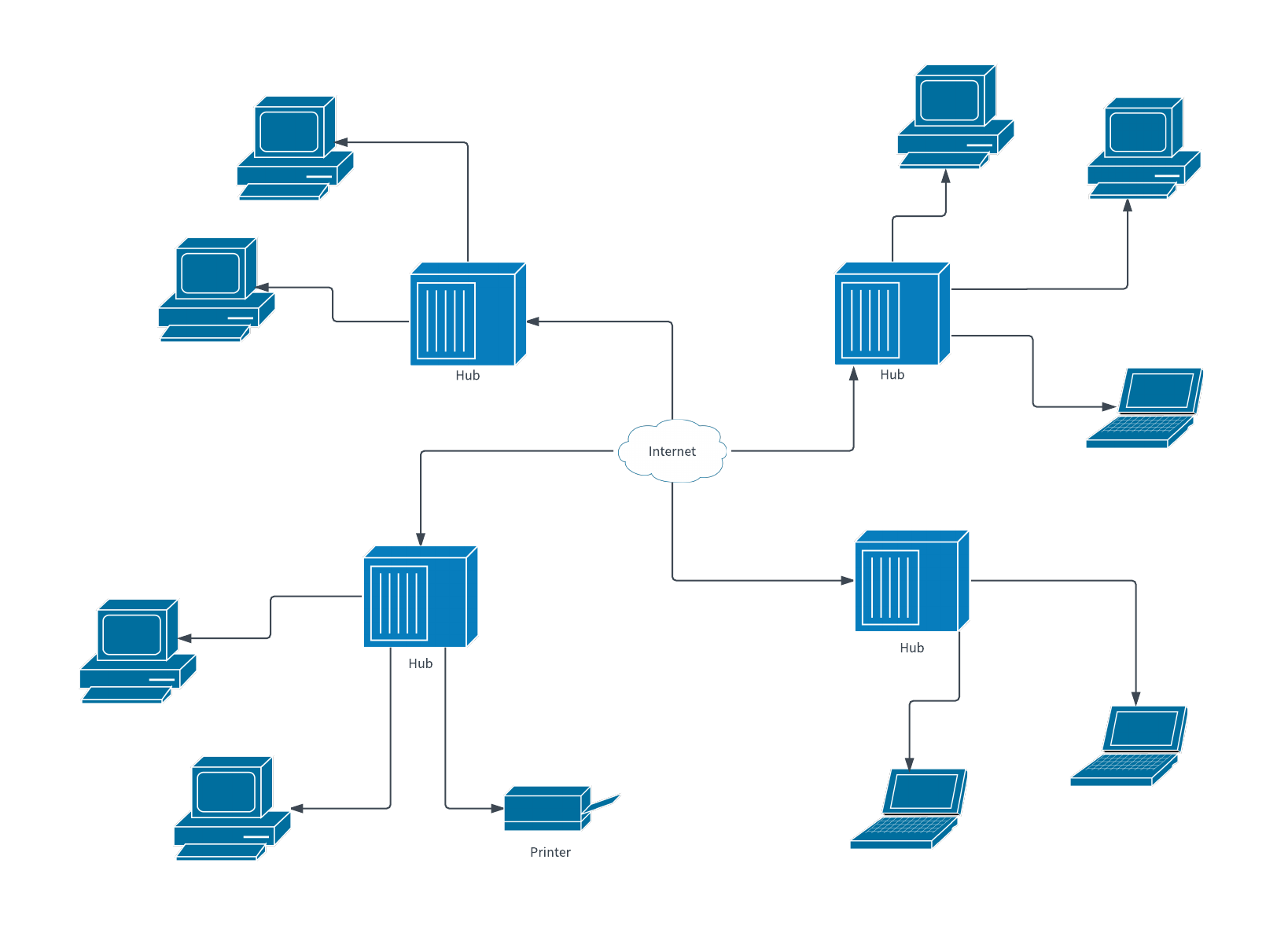


Fig Database Network

This is called a database networking model where the data hub, print, er, and the several systems are connected to the internet. This is how the data is processed and stored in the system to get seethe security of the data and to get management of data effectively.

## Database Infrastructural Diagram

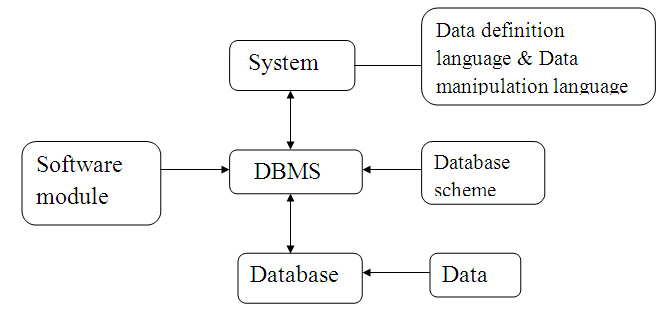


Fig : Database Infrastructure

This is an infrastructure model that is presenting the system with data definition, database scheme, data, database, DBMS, and the software modules.

While creating a database, all these are required effectively where the data in the database system can be recorded effectively. This is how the data is managed and results are achieved better as expected (Dasgupta, et, al, 2020).

## Data Types and schema

## Schemas

Graphical user interface, table

Description automatically generated

Fig : Schemas database

This is called relational schema that creates the tables for the data that make it easy to define all the data into the database in its real form whether integer or variable, dates or time. This is how the instructions are provided for the real outputs. This is required to perform and the above-shown diagram is presenting all the data with the tables that is containing its types effectively.

## Sequencing data

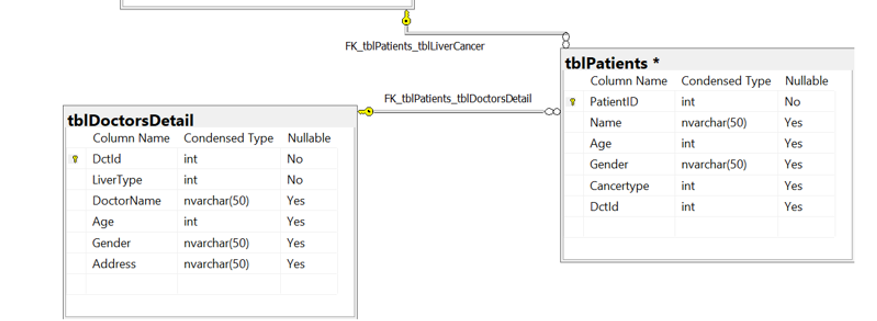


Fig : Sequencing data

## Pseudo Code

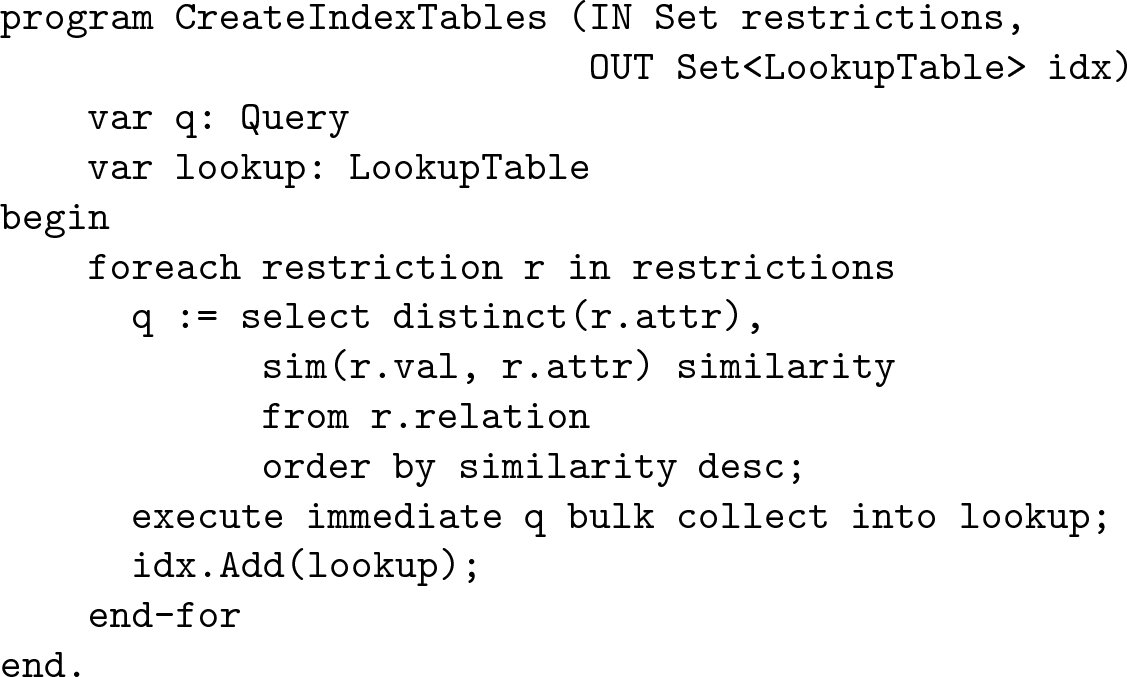


Fig : Pseudo Code for Query Processing

Pseudo codes are the process or performing operations or the syntax of the data in the database to get the outputs. This is given for the query processing in the database SQL. This is showing how to process a query in an SQL database (Liu, et, al, 2020).

## Data Tables

Table

Description automatically generated

# Conclusion

This documentation is based on developing a database system that is effective to handle all the data related to the liver cancer, doctors, and the patients in the system to keep monitoring the cancer and the data effectively. The documentation has shown the significance of the ERD model and the relational schema while creating the database model as these are important to get the data entities and their attributes with their storing pattern to make it easy for the admin. The report has given the appropriate development of the database in the diagram such as the infrastructure diagram and the network diagram. This is how operations can be planned to deploy in the database and the outputs can be received appropriately. The database development is done where several queries are processed effectively in the report.

# References

Chiang, C. J., Wang, Y. W., & Lee, W. C. (2019). Taiwan's nationwide cancer registry system of 40 years: past, present, and future. *J Formos Med Assoc*, *118*(5), 856-858.

Liang, W. C., Wong, C. W., Liang, P. P., Shi, M., Cao, Y., Rao, S. T., ... & Zhang, J. F. (2019). Translation of the circular RNA circβ-catenin promotes liver cancer cell growth through activation of the Wnt pathway. *Genome biology*, *20*(1), 1-12.

Lv, C. X., Zhang, Q., Li, C., Li, Y. G., Li, E. T., Li, Z. R., & Wang, T. C. (2022). Complement Factor H is a Novel Biomarker for Diagnosis and Prognosis of Patients with Liver Cancer. *Indian Journal of Pharmaceutical Sciences*, 48-58.

Cheung, P. K., Ma, M. H., Tse, H. F., Yeung, K. F., Tsang, H. F., Chu, M. K. M., ... & Wong, S. C. C. (2019). The applications of metabolomics in the molecular diagnostics of cancer. *Expert review of molecular diagnostics*, *19*(9), 785-793.

Dasgupta, P., Henshaw, C., Youlden, D. R., Clark, P. J., Aitken, J. F., & Baade, P. D. (2020). Global trends in incidence rates of primary adult liver cancers: A systematic review and meta-analysis. *Frontiers in oncology*, *10*, 171.

Liu, M., Yan, Q., Sun, Y., Nam, Y., Hu, L., Loong, J. H., ... & Guan, X. Y. (2020). A hepatocyte differentiation model reveals two subtypes of liver cancer with different oncofetal properties and therapeutic targets. *Proceedings of the National Academy of Sciences*, *117*(11), 6103-6113.