11253444

Abstract

This report provides a business analysis of TechMart, a hospital information management company. It will introduce a schema for a hospital data management system as well as implement the design in MySQL. Furthermore, matplotlib will be utilised to visualise data that has been extracted from the TechMart database. Moreover, ethical consideration shall be discussed in handling of the said data.



BIg DATA  
A relational data base design with implementaion using an established methodology  
604IT

COURSEWORK 1

Contents

[SECTION:1 INTRODUCTION 5](#_Toc157083729)

[SECTION 2: CRITICAL BUSINESS ANALYSIS 5](#_Toc157083730)

[Identify and describe the key business objectives and requirements of TechMart Inc. 5](#_Toc157083731)

[Company 5](#_Toc157083732)

[Objective 5](#_Toc157083733)

[Analyse and evaluate the existing data management system and identify its shortcomings 5](#_Toc157083734)

[Problem 5](#_Toc157083735)

[Requirement 5](#_Toc157083736)

[Assumptions 5](#_Toc157083737)

[Provide a well-structured and coherent analysis of how the proposed database solution will address the identified business objectives 6](#_Toc157083738)

[Analysis 6](#_Toc157083739)

[SECTION 3: RELATIONAL DATABASE DESIGN 6](#_Toc157083740)

[Design an appropriate relational database schema that effectively captures the required data entities, attributes, and relationships 6](#_Toc157083741)

[Conceptual 6](#_Toc157083742)

[Justify the design decisions, ensuring it adheres to the principles of normalisation and is optimised for data retrieval and management 12](#_Toc157083743)

[Logical design 12](#_Toc157083744)

[Demonstrate a clear understanding of primary keys, foreign keys, and other constraints necessary for maintaining data integrity 13](#_Toc157083745)

[Design 13](#_Toc157083746)

[SECTION 4: DATABASE IMPLEMENTATION 16](#_Toc157083747)

[Implement the designed database schema using a relational database management system (e.g., MySQL, PostgreSQL, or SQL Server 16](#_Toc157083748)

[Development 16](#_Toc157083749)

[Physical 16](#_Toc157083750)

[Create 17](#_Toc157083751)

[Populate the database with sample data representative of TechMart Inc.'s business operations 21](#_Toc157083752)

[Insert 21](#_Toc157083753)

[Provide SQL scripts to create the tables, insert data, and demonstrate the functionality of the implemented database 29](#_Toc157083754)

[Read 30](#_Toc157083755)

[Update 34](#_Toc157083756)

[Delete 36](#_Toc157083757)

[SECTION 5: DATA ANALYSIS AND VISUALISATION 37](#_Toc157083758)

[Apply appropriate data analysis techniques to extract meaningful insights from the TechMart Inc. database 37](#_Toc157083759)

[Implementation 37](#_Toc157083760)

[Interpret the visualisations and identify trends, patterns, and potential areas of improvement for TechMart Inc. 42](#_Toc157083761)

[Utilise data visualisation tools (e.g., Tableau, Power BI, or matplotlib) to create insightful visual representations of the data 43](#_Toc157083762)

[Identifying the number of patients per disease using disease and patient\_disease table 43](#_Toc157083763)

[Counting the number of patients per area using COUNT and JOIN query with address\_code and patient table 44](#_Toc157083764)

[Identifyng the distribution of diseases per city utilising patient, disease, address\_code and patient\_disease table 45](#_Toc157083765)

[Scatter plot utilitsing age, patient, disease and patient\_disease table 46](#_Toc157083766)

[Box plot to query age distribution by disease 47](#_Toc157083767)

[Potential areas of improvement 48](#_Toc157083768)

[SECTION 6: ETHICAL AND LEGAL CONSIDERATIONS 48](#_Toc157083769)

[Discuss the ethical considerations related to handling sensitive data (e.g., customer information, financial data) and propose strategies to ensure data privacy and security 48](#_Toc157083770)

[Address legal requirements, such as data protection laws, and explain how TechMart Inc. can comply with relevant regulations 48](#_Toc157083771)

[SECTION 7: REFERENCES 49](#_Toc157083772)

[Figure 1 Hospital management system ER diagram 7](#_Toc157083773)

[Figure 2 Hospital entity and attributes 7](#_Toc157083774)

[Figure 3 Doctor entity with attributes 8](#_Toc157083775)

[Figure 4 Nurse entity with attributes 8](#_Toc157083776)

[Figure 5 Patient entity with attributes 9](#_Toc157083777)

[Figure 6 Vital signs entity with attributes 9](#_Toc157083778)

[Figure 7 NEWS score entity with attributes 10](#_Toc157083779)

[Figure 8 Lab reports entity with attributes 10](#_Toc157083780)

[Figure 9 Lab name entity and attributes 11](#_Toc157083781)

[Figure 10 Room entity and attributes 11](#_Toc157083782)

[Figure 11 Room type entity and attributes 11](#_Toc157083783)

[Figure 12 Disease entity and attributes 12](#_Toc157083784)

[Figure 13 Identified Entity 12](#_Toc157083785)

[Figure 14 Entity with attributes 12](#_Toc157083786)

[Figure 15 First normal form 13](#_Toc157083787)

[Figure 16 Second normal form with primary keys 13](#_Toc157083788)

[Figure 17 Second normal form with foreign keys 14](https://d.docs.live.net/eb63859381d25af5/11253444_603IT_coursework1_bigdata.docx#_Toc157083789)

[Figure 18 Third normal form 15](https://d.docs.live.net/eb63859381d25af5/11253444_603IT_coursework1_bigdata.docx#_Toc157083790)

[Figure 19 Creating tables in MySQL 16](https://d.docs.live.net/eb63859381d25af5/11253444_603IT_coursework1_bigdata.docx#_Toc157083791)

[Figure 20 Hospital table MySQL code 17](#_Toc157083792)

[Figure 21 Creating doctor 17](#_Toc157083793)

[Figure 22 Creating nurse 17](#_Toc157083794)

[Figure 23 Creating patient 18](#_Toc157083795)

[Figure 24 Creating vital\_signs 18](#_Toc157083796)

[Figure 25 Creating new\_score 18](#_Toc157083797)

[Figure 26 Creating disease 19](#_Toc157083798)

[Figure 27 Creating patient\_disease 19](#_Toc157083799)

[Figure 28 Creating address\_code 19](#_Toc157083800)

[Figure 29 Creating lab\_name 19](#_Toc157083801)

[Figure 30 Creating lab\_report 20](#_Toc157083802)

[Figure 31 Creating lab\_report\_name 20](#_Toc157083803)

[Figure 32 Creating patient\_lab\_report 20](#_Toc157083804)

[Figure 33 Creating patient\_vital\_signs 20](#_Toc157083805)

[Figure 34 Create room\_type 21](#_Toc157083806)

[Figure 35 Create room 21](#_Toc157083807)

[Figure 36 Inserting a new patient 21](https://d.docs.live.net/eb63859381d25af5/11253444_603IT_coursework1_bigdata.docx#_Toc157083808)

[Figure 37 Inserting patients in the table 22](#_Toc157083809)

[Figure 38 Output 22](#_Toc157083810)

[Figure 39 Inserting in address\_code 23](#_Toc157083811)

[Figure 40 Inserting in disease 23](#_Toc157083812)

[Figure 41 Inserting in doctor 23](#_Toc157083813)

[Figure 42 Inserting in hospital 24](#_Toc157083814)

[Figure 43 Inserting in lab\_name 24](#_Toc157083815)

[Figure 44 Inserting in lab\_report 24](#_Toc157083816)

[Figure 45 Inserting in lab\_report\_name 25](#_Toc157083817)

[Figure 46 Inserting in news\_score 25](#_Toc157083818)

[Figure 47 Inserting in nurse 25](#_Toc157083819)

[Figure 48 Inserting in patient\_disease 26](#_Toc157083820)

[Figure 49 Inserting in patient\_lab\_report 26](#_Toc157083821)

[Figure 50 Inserting in patient\_vital\_signs 26](#_Toc157083822)

[Figure 51 Inserting in room 27](#_Toc157083823)

[Figure 52 Inserting in room\_type 27](#_Toc157083824)

[Figure 53 Inserting in vital\_signs 28](#_Toc157083825)

[Figure 54 Inserting record in vital signs and NEWS\_score 29](https://d.docs.live.net/eb63859381d25af5/11253444_603IT_coursework1_bigdata.docx#_Toc157083826)

[Figure 55 Viewing a patient record 30](#_Toc157083827)

[Figure 56 Counting gender 30](#_Toc157083828)

[Figure 57 Patient and all their diseases 31](#_Toc157083829)

[Figure 58 Patient joined with doctor name 32](#_Toc157083830)

[Figure 59 Total number of patients per address code 33](#_Toc157083831)

[Figure 60 Maximum age of patient in an address 33](#_Toc157083832)

[Figure 61 Updating a patient record 34](#_Toc157083833)

[Figure 62 Vital signs converted to news score 35](https://d.docs.live.net/eb63859381d25af5/11253444_603IT_coursework1_bigdata.docx#_Toc157083834)

[Figure 63 Deleting row 3 36](#_Toc157083835)

[Figure 64 Installing mysql-connector-python and connecting to Mysql database 37](#_Toc157083836)

[Figure 65 Patient data retrieved from company database in mysql 38](#_Toc157083837)

[Figure 66 Converting table to csv file 39](#_Toc157083838)

[Figure 67 Conversion output 40](#_Toc157083839)

[Figure 68 Histogram of patient admission by age 41](#_Toc157083840)

[Figure 69 Bar chart comparing admission by gender 42](#_Toc157083841)

[Figure 70 Query:Counting the number of patients per disease 43](#_Toc157083842)

[Figure 71 Query: Count the number of patients in an area 44](https://d.docs.live.net/eb63859381d25af5/11253444_603IT_coursework1_bigdata.docx#_Toc157083843)

[Figure 72 Query: Disease distribution in each city 45](https://d.docs.live.net/eb63859381d25af5/11253444_603IT_coursework1_bigdata.docx#_Toc157083844)

[Figure 73 Query: age distribution of patients with disease\_id 1 46](https://d.docs.live.net/eb63859381d25af5/11253444_603IT_coursework1_bigdata.docx#_Toc157083845)

[Figure 74 Query: age distribution across diseases 47](https://d.docs.live.net/eb63859381d25af5/11253444_603IT_coursework1_bigdata.docx#_Toc157083846)

# SECTION:1 INTRODUCTION

This report aims to critically analyse the business requirements of a fictional health information company, show how I designed and developed a relational database, and apply relevant data visualisation techniques to gain insight from the data.

# SECTION 2: CRITICAL BUSINESS ANALYSIS

## Identify and describe the key business objectives and requirements of TechMart Inc.

### Company

TechMart is as a premier Health Information Management (HIM) company.

### Objective

Dedicated to improve efficiency in the health sector through database management system development. The system will enhance workflow, optimise resource allocation and improve patient’s record management.

## Analyse and evaluate the existing data management system and identify its shortcomings

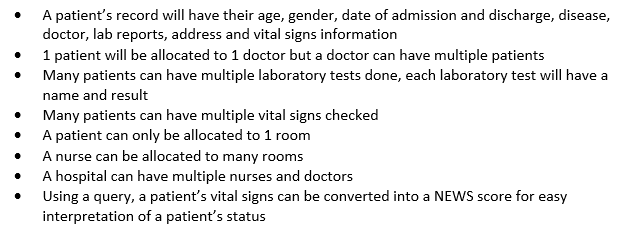
### Problem

There are multiple database systems existing within different hospitals with no consistent format. It has been identified that it is causing duplication of information which results in resource wastage.

### Requirement

The company aims to develop a robust hospital database management system that can be used uniformly by different hospitals for interoperability. The system will consist of patient, patient’s vital signs, doctors, nurses, lab reports, hospital and room records.

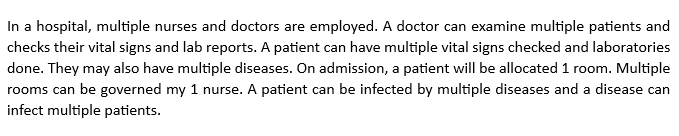
### Assumptions



## Provide a well-structured and coherent analysis of how the proposed database solution will address the identified business objectives

### Analysis

In a top-down view the entities interact this way:



# SECTION 3: RELATIONAL DATABASE DESIGN

## Design an appropriate relational database schema that effectively captures the required data entities, attributes, and relationships

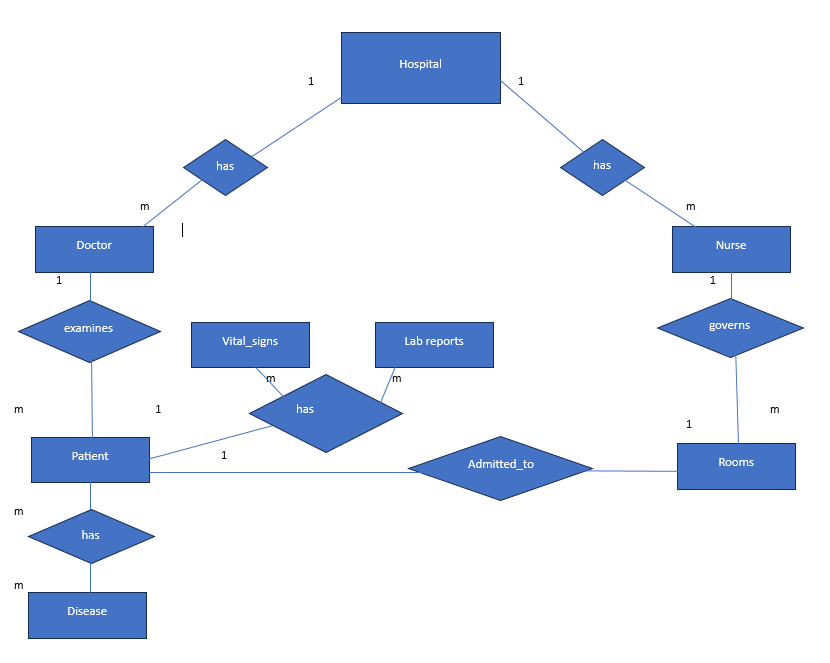
### Conceptual

Entity identification: disease, patient, vital signs, doctor, lab report, rooms, nurse and hospital.

#### ER diagram

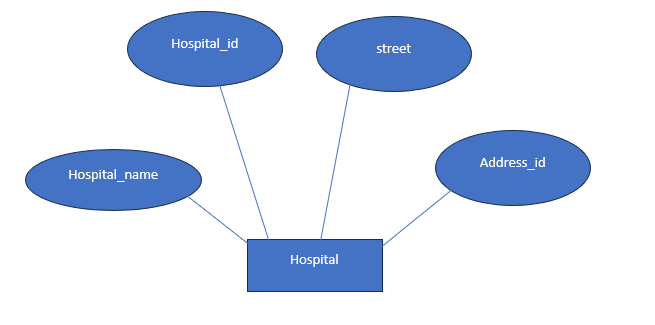
Conceptually, Figure 1 is the ER diagram of the entities. The entities and their attribute have bene placed in a different image.

Figure 1 Hospital management system ER diagram



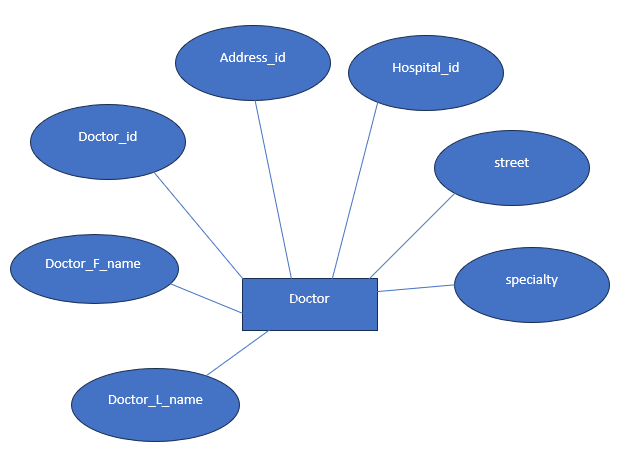
(Image supplied by author)

Figure 2 Hospital entity and attributes



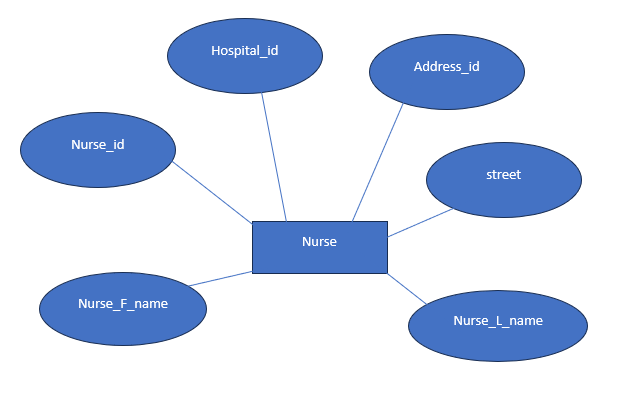
(Image supplied by author)

Figure 3 Doctor entity with attributes



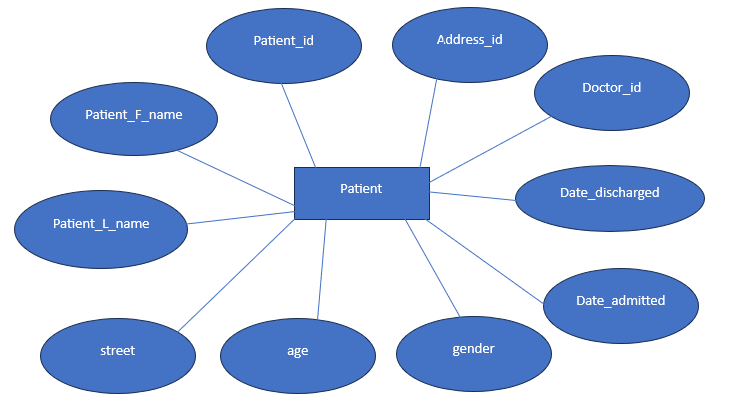
(Image supplied by author)

Figure 4 Nurse entity with attributes



(Image supplied by author)

Figure 5 Patient entity with attributes



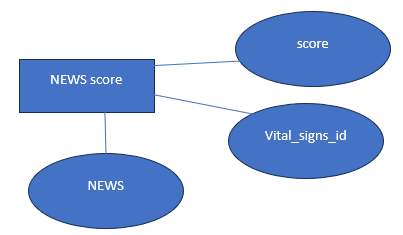
(Image supplied by author)

Figure 6 Vital signs entity with attributes



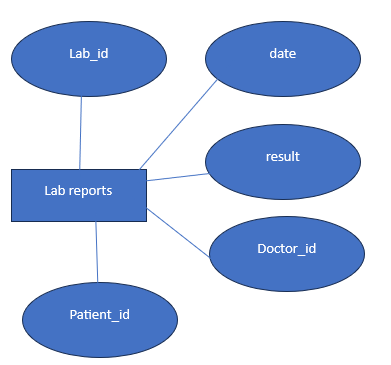
(Image supplied by author)

Figure 7 NEWS score entity with attributes



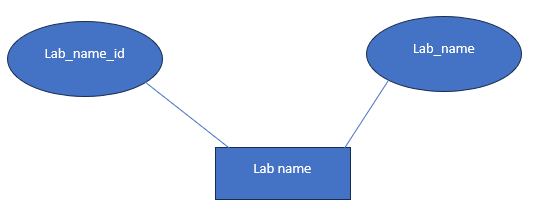
(Image supplied by author)

Figure 8 Lab reports entity with attributes



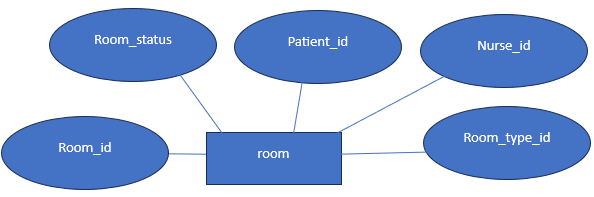
(Image supplied by author)

Figure 9 Lab name entity and attributes



(Image supplied by author)

Figure 10 Room entity and attributes



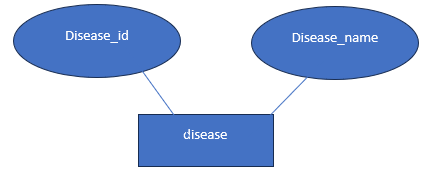
(Image supplied by author)

Figure 11 Room type entity and attributes



(Image supplied by author)

Figure 12 Disease entity and attributes



(Image supplied by author)

## Justify the design decisions, ensuring it adheres to the principles of normalisation and is optimised for data retrieval and management

### Logical design

Logically, my initial non-normalised table would look like Figure 13:

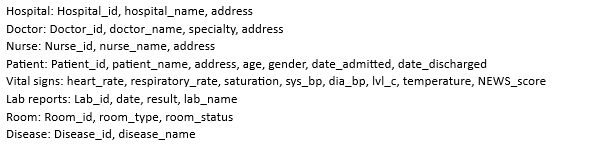
Figure 13 Identified Entity



(Image supplied by author)

After identifying the entities of the hospital management system, I proceeded to enumerate the attributes that is associated with each entity.

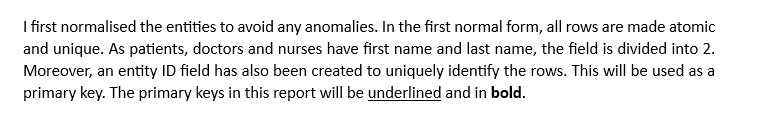
Figure 14 Entity with attributes



(Image supplied by author)

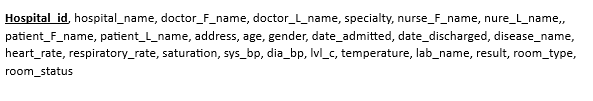
## Demonstrate a clear understanding of primary keys, foreign keys, and other constraints necessary for maintaining data integrity

### Design



#### 1st normal form

Figure 15 First normal form



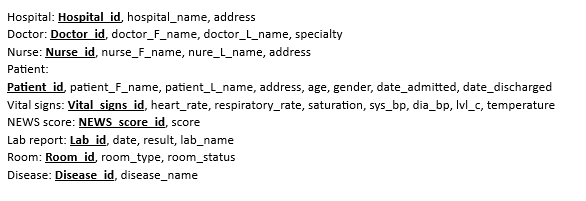
(Image supplied by author)

#### 2nd normal form

In the second form of normalisation, there are no partial dependencies. I separated the fields that are not functionally dependent to the **Hospital\_id**. As doctor, nurse, patient, vital signs, lab reports, rooms and disease are separate entities, I placed them on different tables along with their attributes.

Moreover, NEWS score is made into a separate entity as it relies on the vital sign’s attributes. NEWS score or National Early Warning score is a scoring system that assigns values depending on the physiological measurements such as oxygen saturation, blood pressure and heart rate. This is then computed and used to identify acutely ill patients (NICE, 2020).

Figure 16 Second normal form with primary keys



(Image supplied by author)

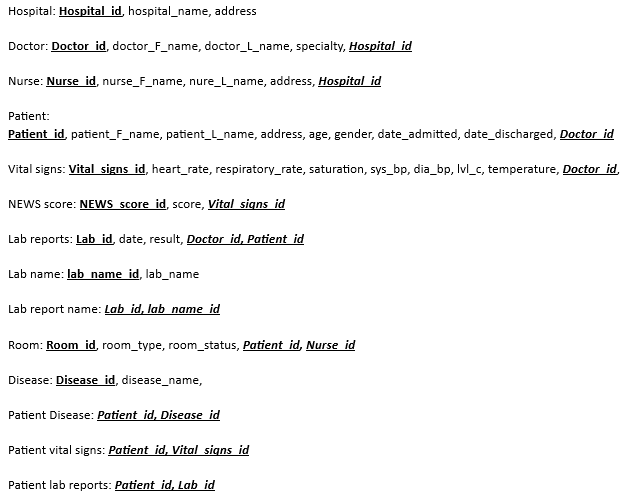
#### Foreign key application

In this stage I have attached foreign keys to the entities. Foreign key is a column that relates to a primary key in another table. In a one is to many relationships; I placed the foreign key of the “one” in to the “many”. In this way the record in the “many” will be associated with a specific “one”.

In a one-to-one relationship such as the patient and room, I made the patient a foreign key in the room table. This is because they may or may not occupy a room and can be discharged without utilising it.

Furthermore, in a many to many relationship such as patient/disease, patient/vital signs, patient/laboratory report and laboratory report/ laboratory name, I have made a separate table. This associative table will break the many to many relationship into 2 one to many relationships.

Figure 17 Second normal form with foreign keys



(Image supplied by author)

#### 3rd normal form

The third form of normalisation satisfies the requirements of the second normal form as well as eliminate transitive functional dependencies. Hospitals, doctors, nurses and patients have addresses and a post code can be used to determine the address of a street, therefore I created another table named “Address code”

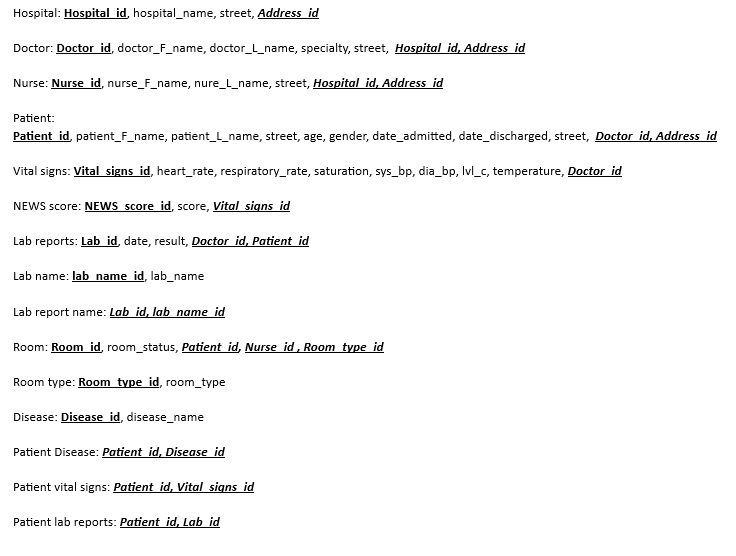


Figure 18 Third normal form

(Image supplied by author)

# SECTION 4: DATABASE IMPLEMENTATION

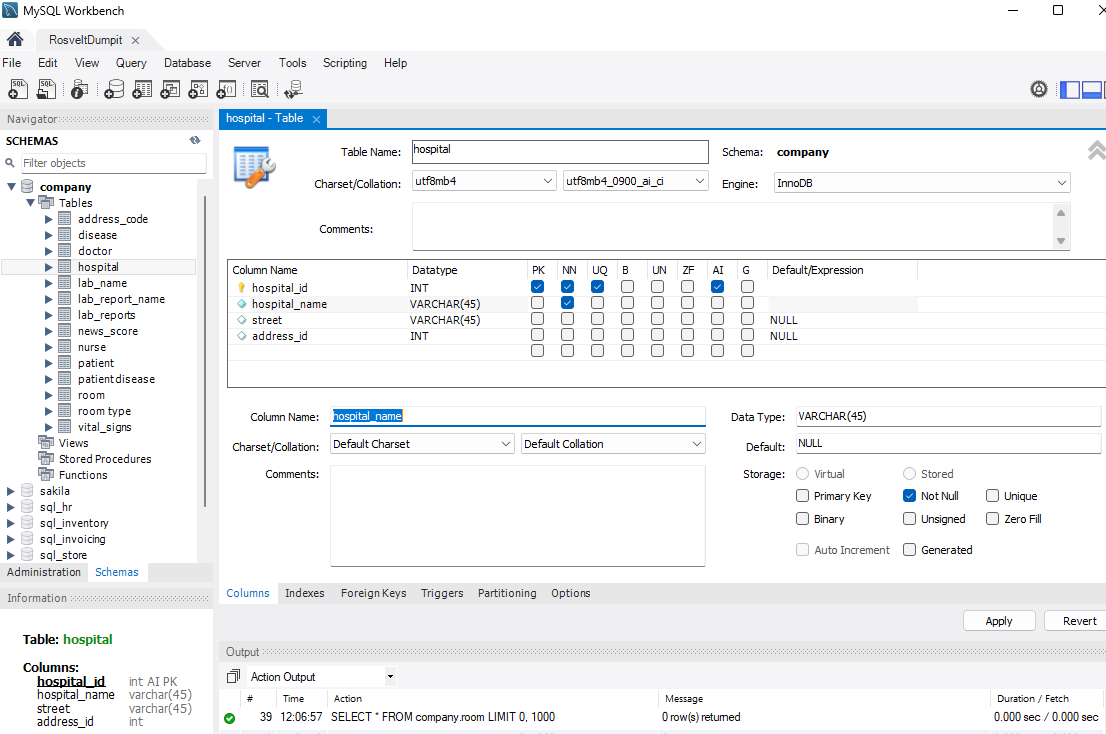
## Implement the designed database schema using a relational database management system (e.g., MySQL, PostgreSQL, or SQL Server

### Development

### Physical

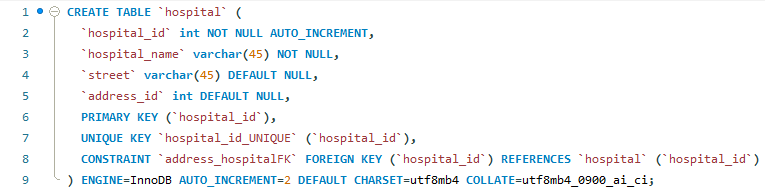
I made use of MySQL to recreate the schema I have made in a relational database management system. Figure 19 illustrates the creation of the hospital table. **Hospital\_id** is used as the primary key, it can not be null, is unique and I have specified that it will autoincrement. Figure 20 provides the code for MySQL.

Figure 19 Creating tables in MySQL



(Image supplied by author)

Figure 20 Hospital table MySQL code

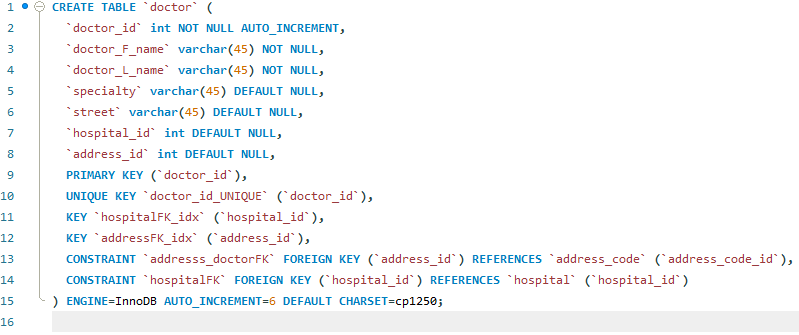


(Image supplied by author)

### Create

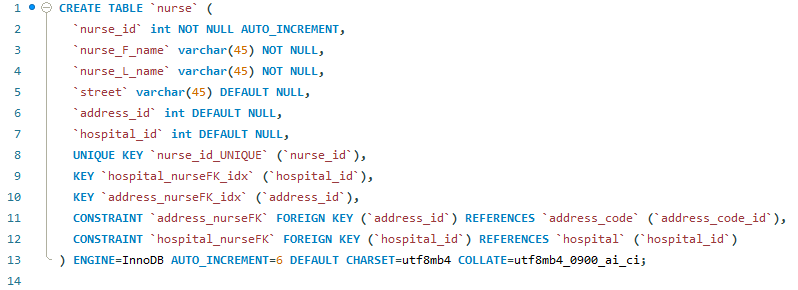
Figure 21 to 27 uses the statement CREATE to make tables.

Figure 21 Creating doctor



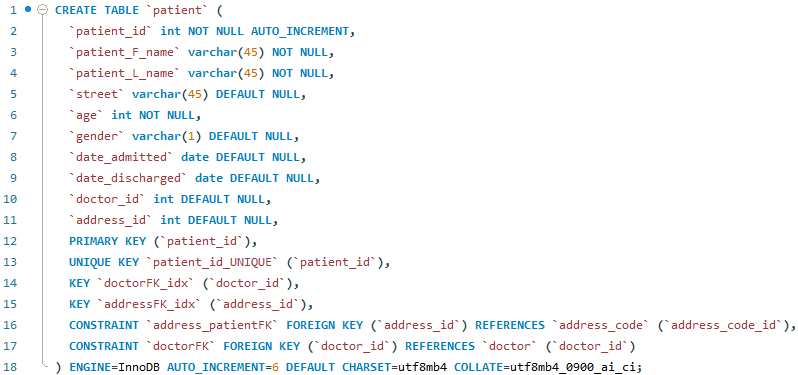
(Image supplied by author)

Figure 22 Creating nurse



(Image supplied by author)

Figure 23 Creating patient



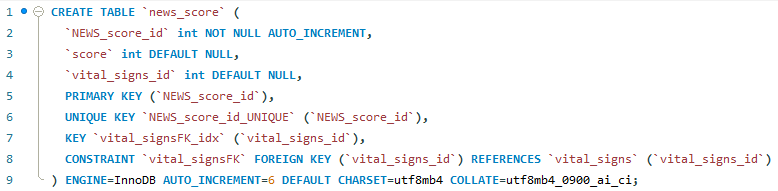
(Image supplied by author)

Figure 24 Creating vital\_signs



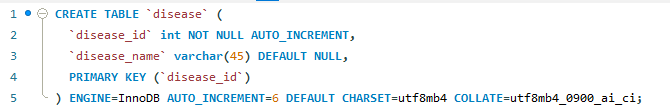
(Image supplied by author)

Figure 25 Creating new\_score



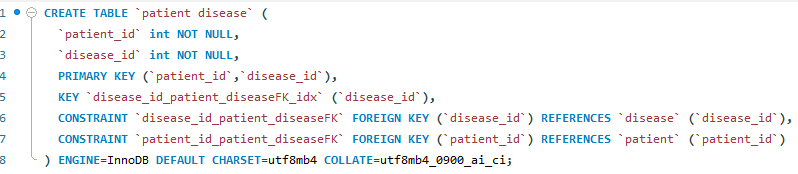
(Image supplied by author)

Figure 26 Creating disease



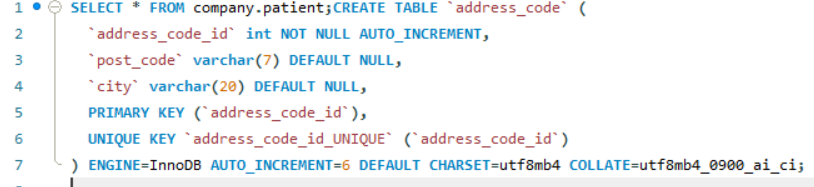
(Image supplied by author)

Figure 27 Creating patient\_disease



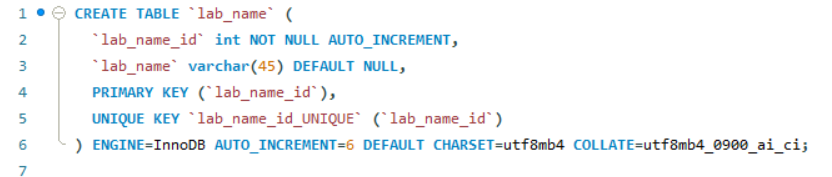
(Image supplied by author)

Figure 28 Creating address\_code



(Image supplied by author)

Figure 29 Creating lab\_name



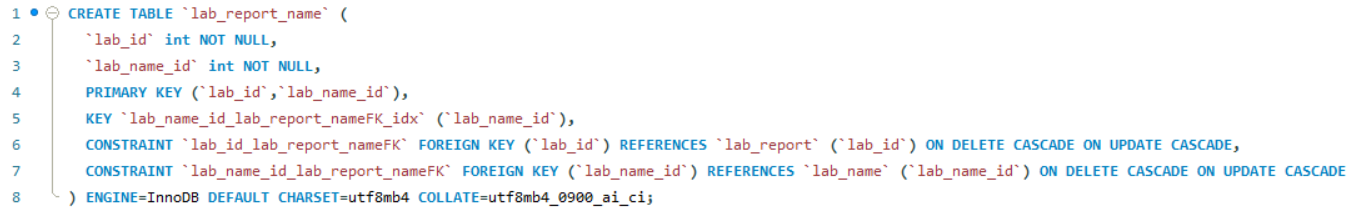
(Image supplied by author)

Figure 30 Creating lab\_report



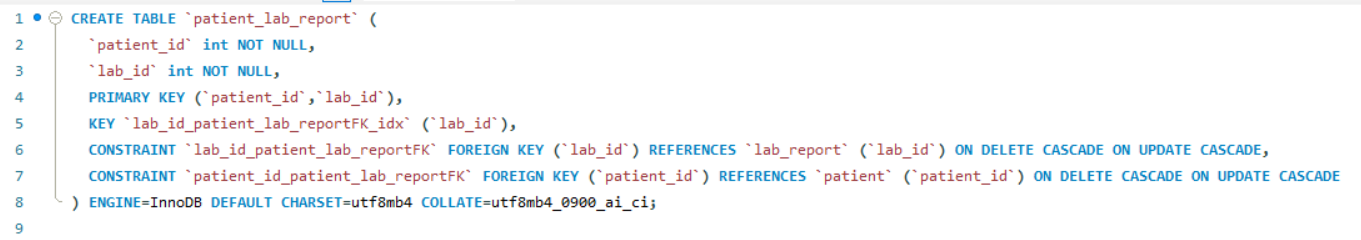
(Image supplied by author)

Figure 31 Creating lab\_report\_name



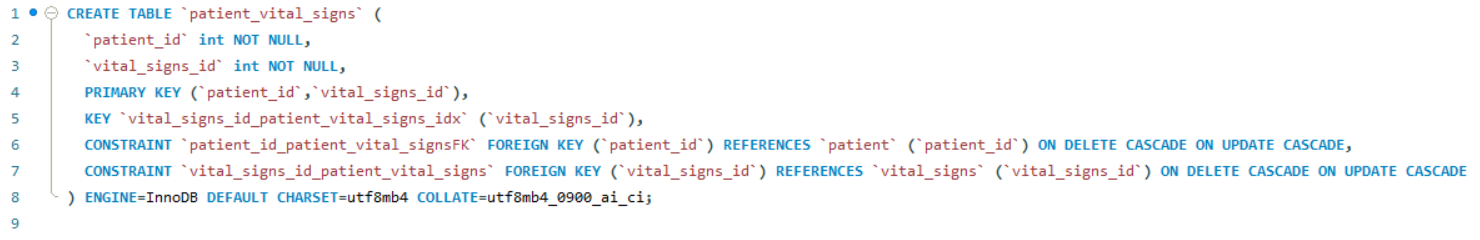
(Image supplied by author)

Figure 32 Creating patient\_lab\_report



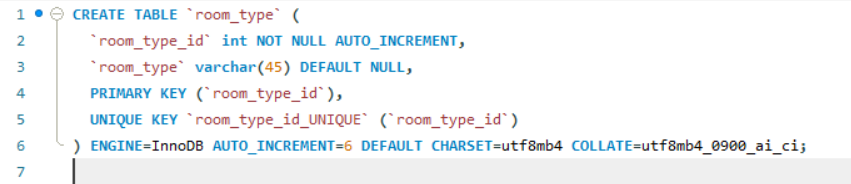
(Image supplied by author)

Figure 33 Creating patient\_vital\_signs



(Image supplied by author)

Figure 34 Create room\_type



(Image supplied by author)

Figure 35 Create room

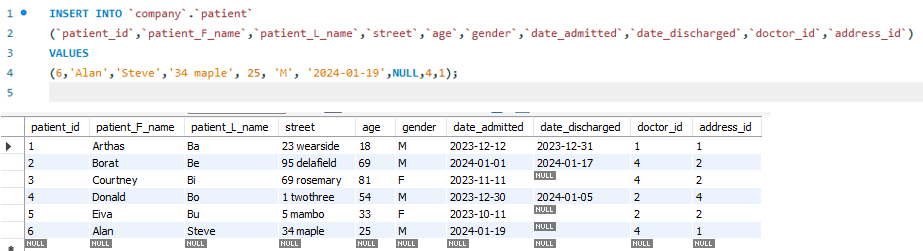


(Image supplied by author)

## Populate the database with sample data representative of TechMart Inc.'s business operations

### Insert

Figure 36 shows how the data is inputted in the patient table. Here I have a 6th patient, a 25-year-old male from London, called Alan Steve.



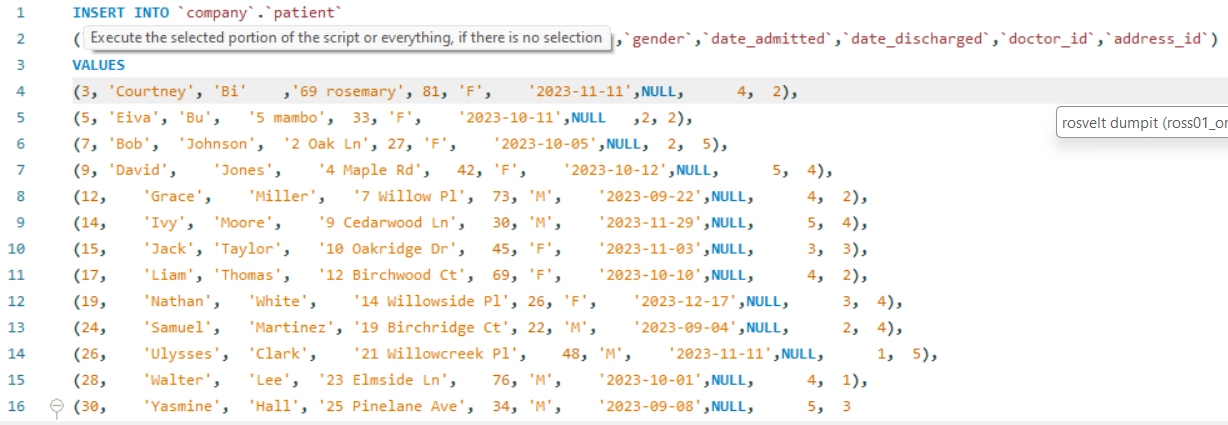
Figure

Figure 36 Inserting a new patient

(Image supplied by author)

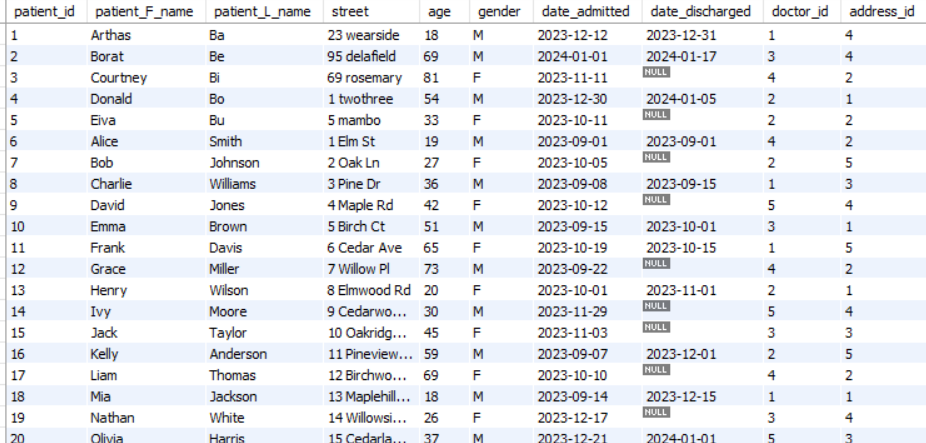
Figure 37 – 53 illustrates the codes used to fill the tables.

Figure 37 Inserting patients in the table



(Image supplied by author)

Figure 38 Output



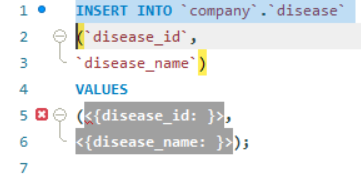
(Image supplied by author)

Figure 39 Inserting in address\_code



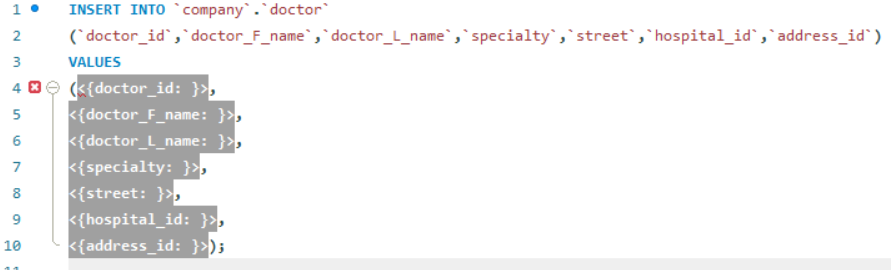
(Image supplied by author)

Figure 40 Inserting in disease



(Image supplied by author)

Figure 41 Inserting in doctor



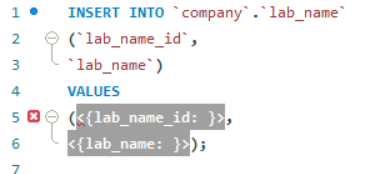
(Image supplied by author)

Figure 42 Inserting in hospital



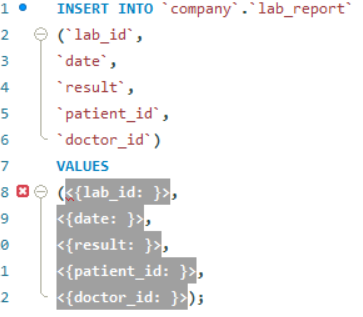
(Image supplied by author)

Figure 43 Inserting in lab\_name



(Image supplied by author)

Figure 44 Inserting in lab\_report



(Image supplied by author)

Figure 45 Inserting in lab\_report\_name



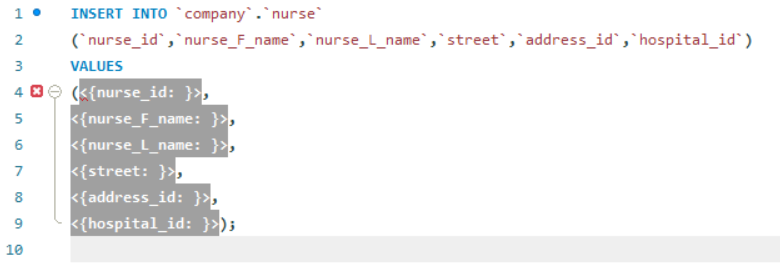
(Image supplied by author)

Figure 46 Inserting in news\_score



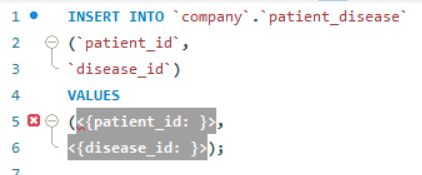
(Image supplied by author)

Figure 47 Inserting in nurse



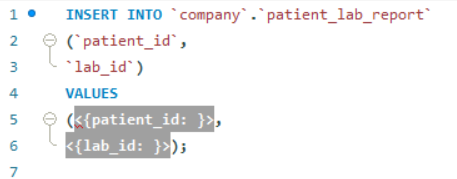
(Image supplied by author)

Figure 48 Inserting in patient\_disease



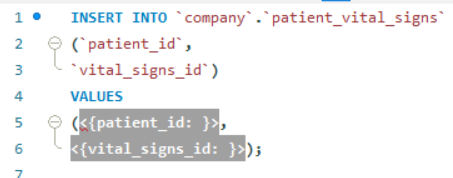
(Image supplied by author)

Figure 49 Inserting in patient\_lab\_report



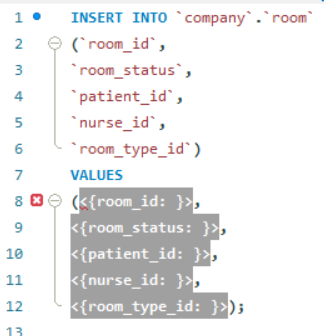
(Image supplied by author)

Figure 50 Inserting in patient\_vital\_signs



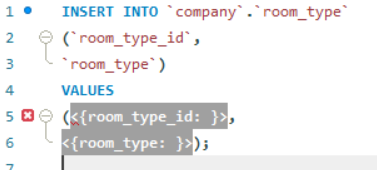
(Image supplied by author)

Figure 51 Inserting in room



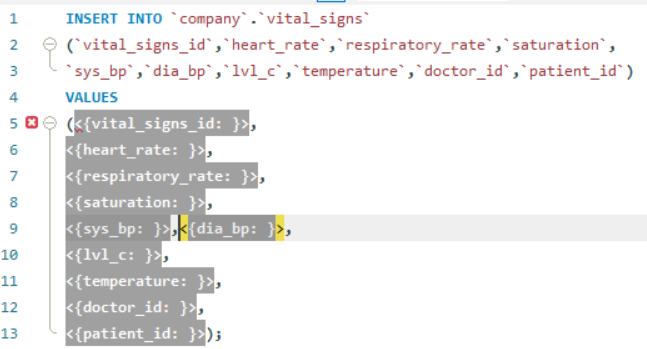
(Image supplied by author)

Figure 52 Inserting in room\_type



(Image supplied by author)

Figure 53 Inserting in vital\_signs



(Image supplied by author)

## Provide SQL scripts to create the tables, insert data, and demonstrate the functionality of the implemented database

Figure 54 shows an example of inserting the vital signs and news core. I have made the score field as NULL as this will be filled when we an UPDATE query is done.

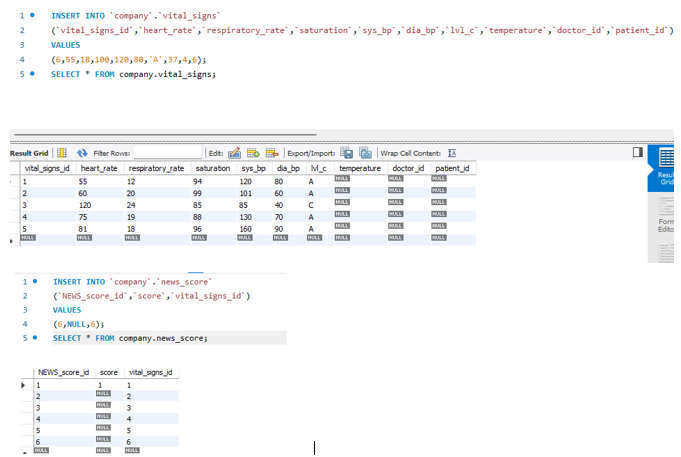


Figure 54 Inserting record in vital signs and NEWS\_score

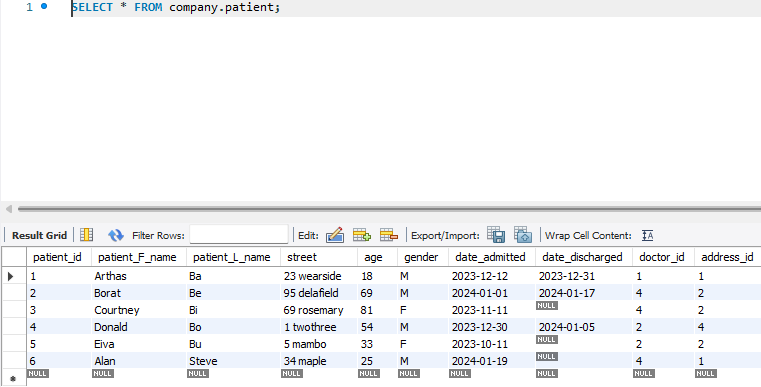
(Image supplied by author)

Figure 54

### Read

Using the SELECT statement, I viewed the existing patient records as seen in Figure 55.

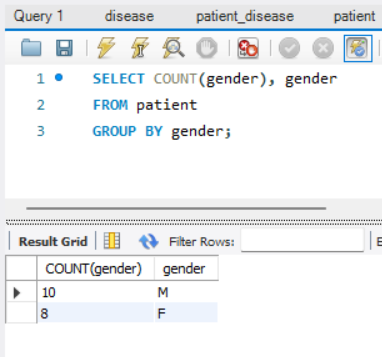
Figure 55 Viewing a patient record



(Image supplied by author)

Figure 56 shows the code to count the number of patients separated by gender.

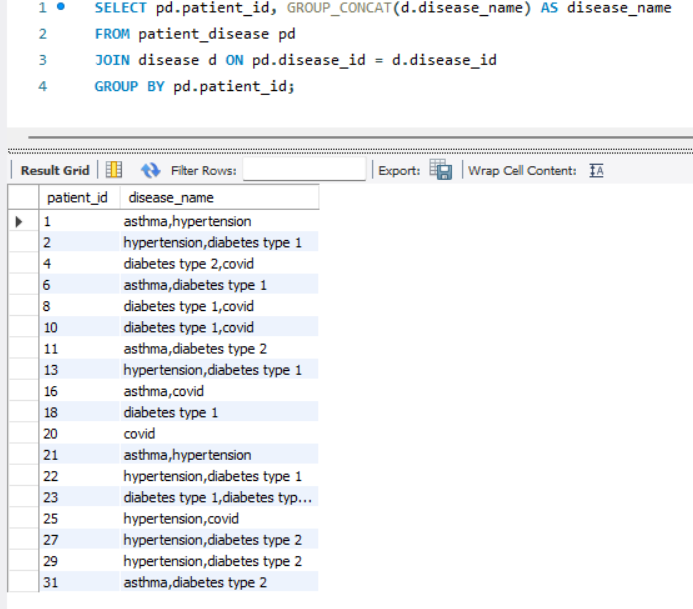
Figure 56 Counting gender



(Image supplied by author)

Figure 57 illustrates how to group the diseases each patient has with the use of JOIN and GROUP BY.

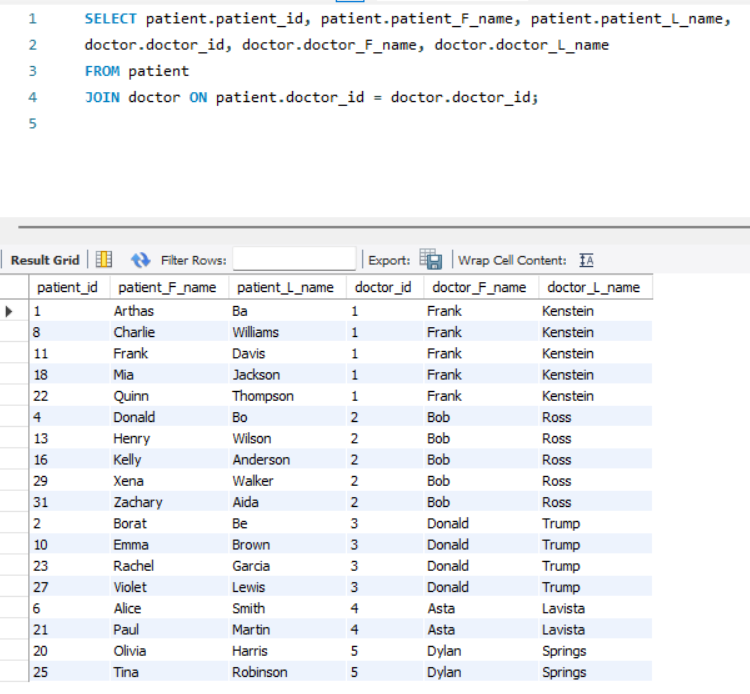
Figure 57 Patient and all their diseases



(Image supplied by author)

Figure 58 illustrates how to join patient and doctor tables.

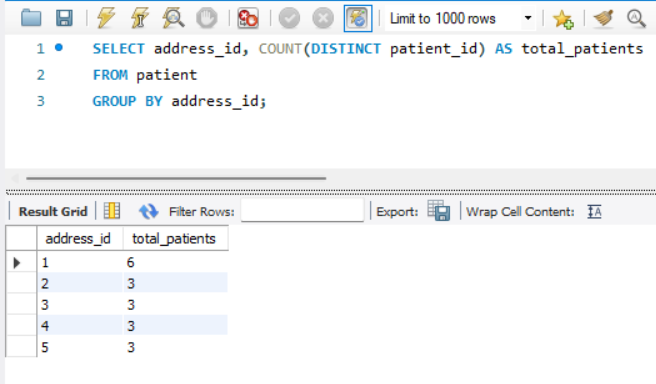
Figure 58 Patient joined with doctor name



(Image supplied by author)

Figure 59 counts the patients by address\_id.

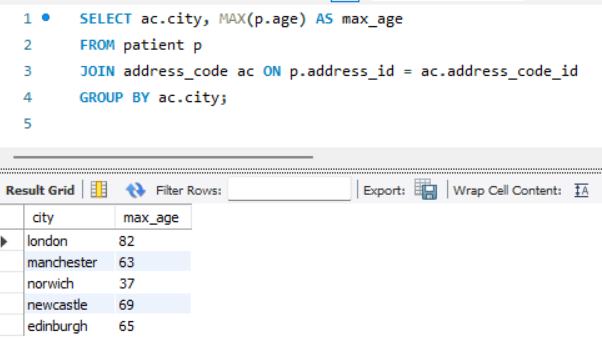
Figure 59 Total number of patients per address code



(Image supplied by author)

Figure 60 shows how the max age of patients by city.

Figure 60 Maximum age of patient in an address

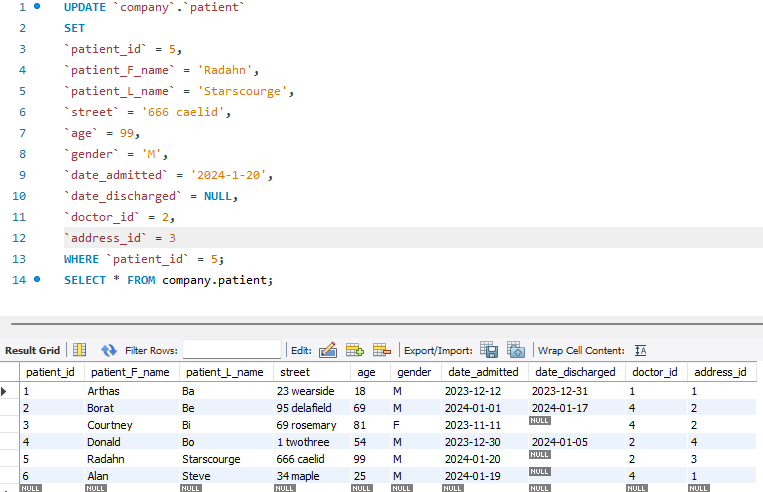


(Image supplied by author)

### Update

Using the UPDATE statement, I modified patient\_id 5 to a different record as seen in Figure 61.

Figure 61 Updating a patient record



(Image supplied by author)

In Figure 62 I used UPDATE to modify the news\_score “score” of patient\_id 6.



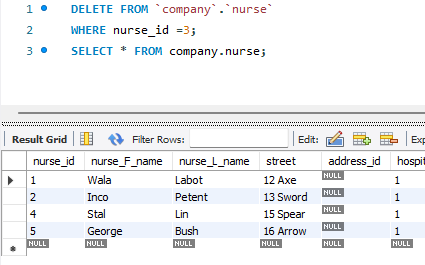
(Image supplied by author)

Figure 62 Vital signs converted to news score

### Delete

Using DELETE I can remove a record in the data base as seen in Figure 63.

Figure 63 Deleting row 3



(Image supplied by author)

# SECTION 5: DATA ANALYSIS AND VISUALISATION

## Apply appropriate data analysis techniques to extract meaningful insights from the TechMart Inc. database

### Implementation

In this stage I used matplotlib in jupyter to represent the data form the TechMart Inc.

Figure 64 Installing mysql-connector-python and connecting to Mysql database

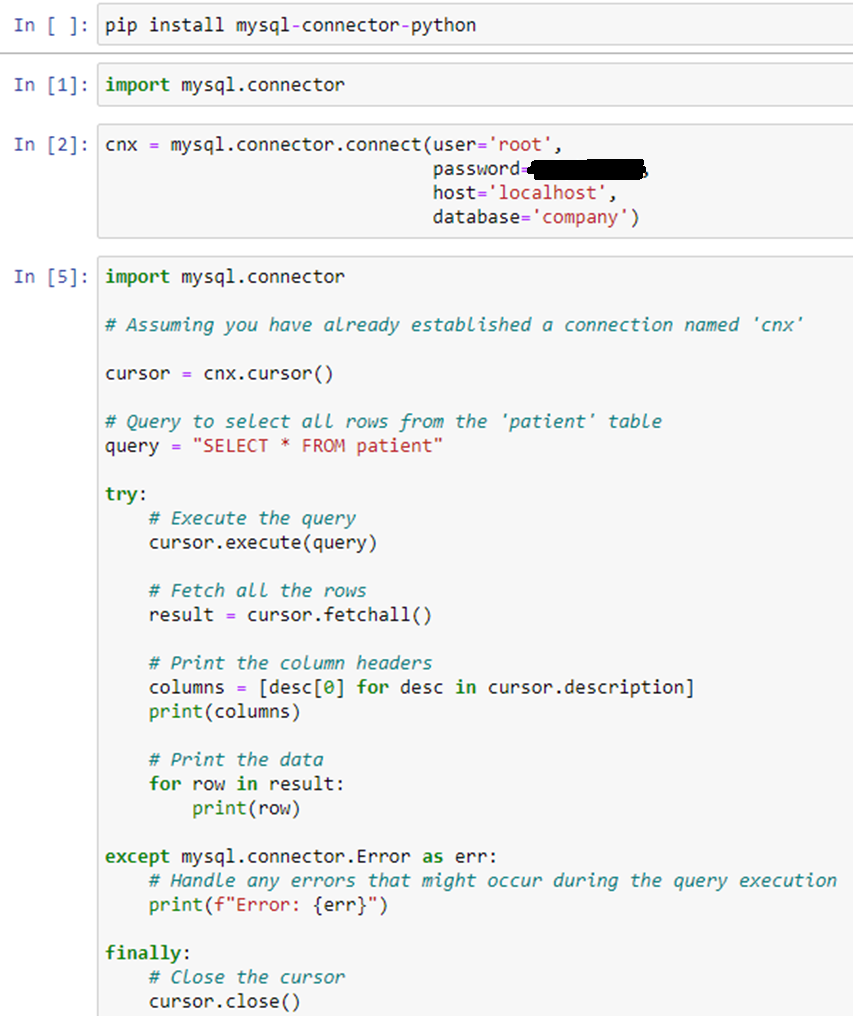


Figure 65 Patient data retrieved from company database in mysql

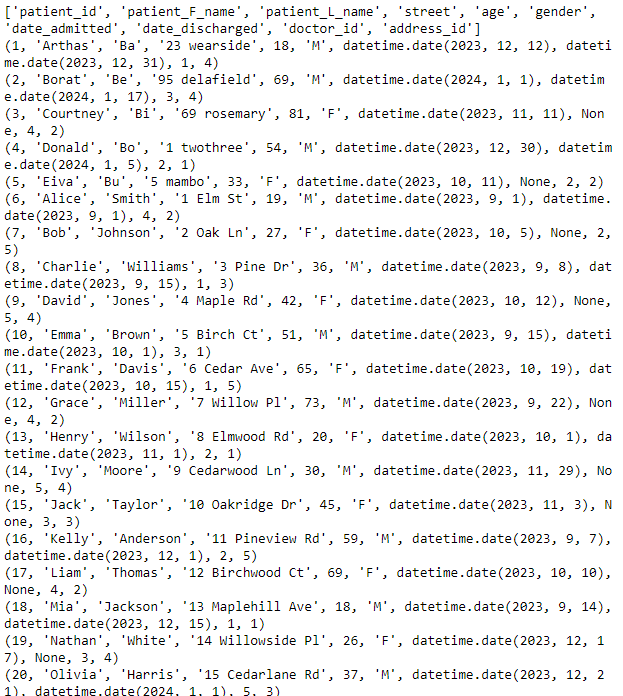


Figure 66 Converting table to csv file

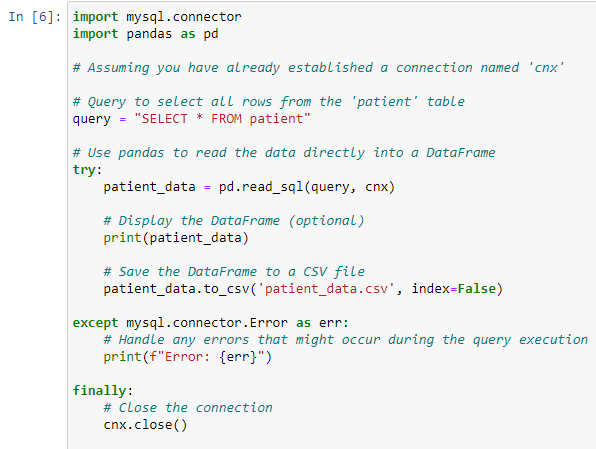


Figure 67 Conversion output

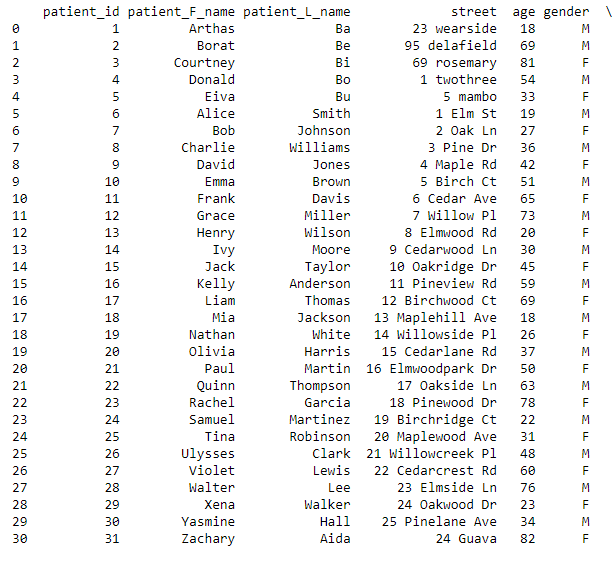
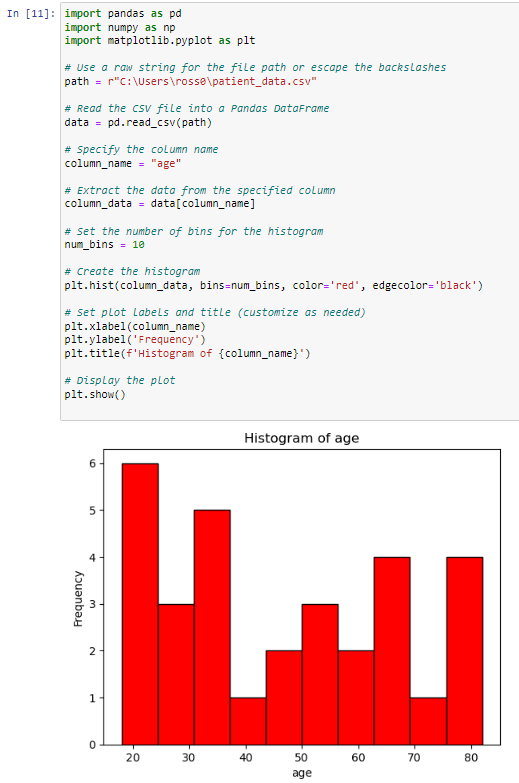


Figure 68 displays a histogram illustrating the distribution of admitted patients by age. It can be gleaned that patients who are in their 20’s are admitted more compared to those in their 40s and 70s.

Figure 68 Histogram of patient admission by age



(Image supplied by author)

## Interpret the visualisations and identify trends, patterns, and potential areas of improvement for TechMart Inc.

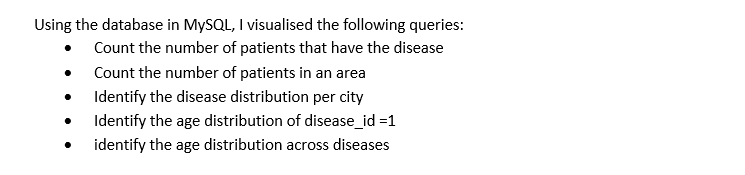
Figure 69 illustrates a slightly higher number of admissions for male patients compared to the female counterpart.

Figure 69 Bar chart comparing admission by gender

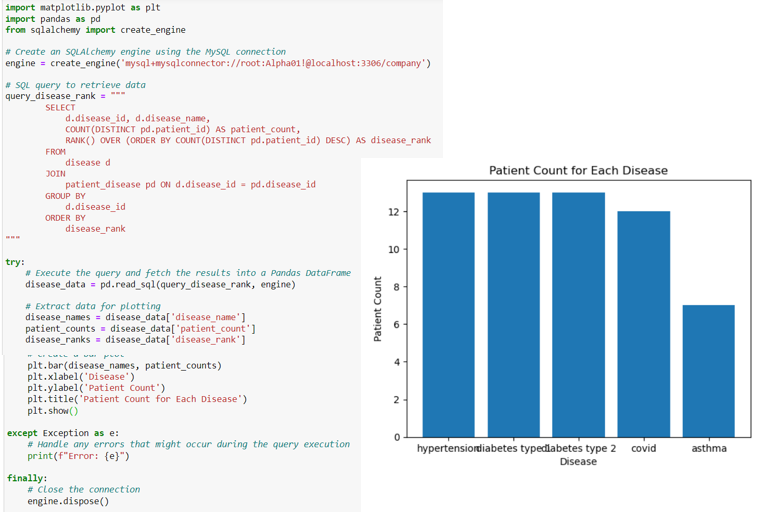


(Image supplied by author)

## Utilise data visualisation tools (e.g., Tableau, Power BI, or matplotlib) to create insightful visual representations of the data



### Identifying the number of patients per disease using disease and patient\_disease table

Figure 70 Query:Counting the number of patients per disease

(Image supplied by author)

### Counting the number of patients per area using COUNT and JOIN query with address\_code and patient table

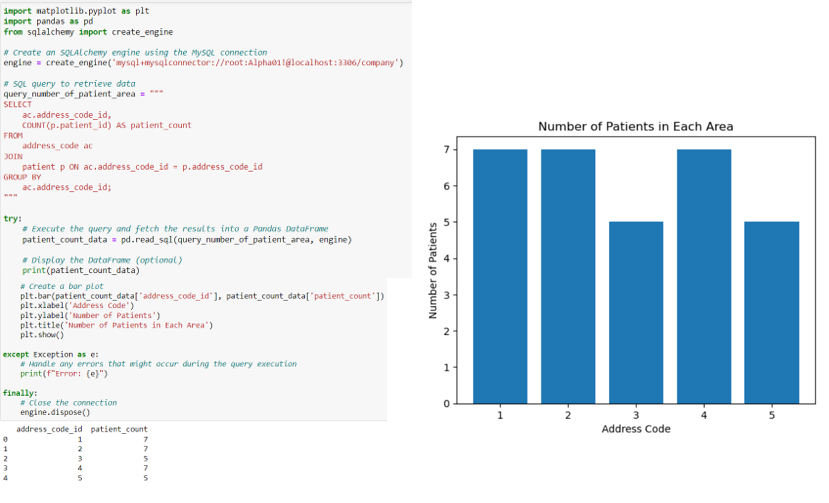


Figure 71 Query: Count the number of patients in an area

(Image supplied by author)

### Identifyng the distribution of diseases per city utilising patient, disease, address\_code and patient\_disease table



Figure 72 Query: Disease distribution in each city

(Image supplied by author)

### Scatter plot utilitsing age, patient, disease and patient\_disease table

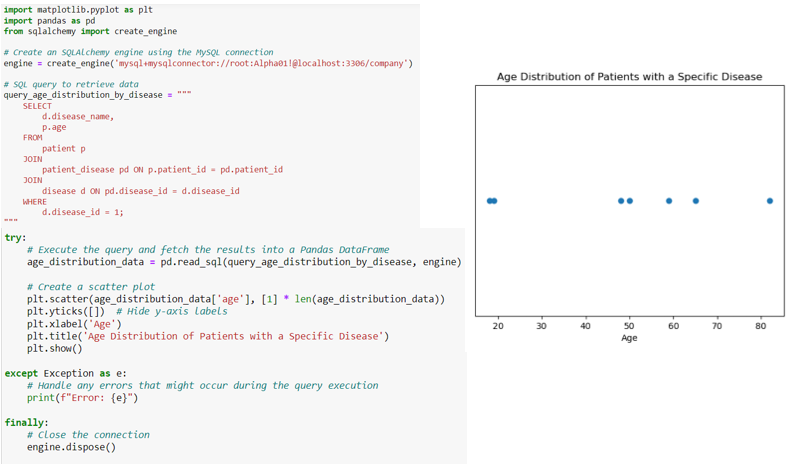


Figure 73 Query: age distribution of patients with disease\_id 1

(Image supplied by author)

### Box plot to query age distribution by disease



Figure 74 Query: age distribution across diseases

(Image supplied by author)

### Potential areas of improvement

TechMart Inc. can further enhance this product by creating a separate entity for outpatients. Currently this product records every patient coming into the hospital but does not count the patients who went home after being seen by a doctor.

# SECTION 6: ETHICAL AND LEGAL CONSIDERATIONS

## Discuss the ethical considerations related to handling sensitive data (e.g., customer information, financial data) and propose strategies to ensure data privacy and security

## Address legal requirements, such as data protection laws, and explain how TechMart Inc. can comply with relevant regulations



# SECTION 7: REFERENCES

NICE. (2020, February 18). *The technology | National Early Warning Score systems that alert to deteriorating adult patients in hospital  | Advice | NICE*. https://www.nice.org.uk/advice/mib205/chapter/The-technology

*Section 3: The Data Protection Act 1998 - NHS Digital*. (n.d.-b). NHS Digital. <https://digital.nhs.uk/data-and-information/looking-after-information/data-security-> and-information-governance/codes-of-practice-for-handling-information-in-health-and- care/a-guide-to-confidentiality-in-health-and-social-care/hscic-guide-to-confidentiality- references/section- 3#:~:text=The%20right%20to%20object%20to,could%20be%20considered%20unfair%20proc essing.