D205 - TGM3 TASK 1: DATA ACQUISITION

Brandon Gillins

Western Governors University

D205 - Task 1

April 27, 2024

Section A:

Question: How many patients with diabetes are admitted to emergency services?

1. Justification and Approach

To address the prevalence of emergency admissions among diabetic patients, an integrated data analysis approach is employed, utilizing three primary data sources within the healthcare database. These include the "Patient" table, providing essential demographic data; the "Health Informatics" table, which records patients' diabetes status; and the "Admission" table, detailing patient admission types.

The analysis begins by identifying emergency admissions through the initial_admissions field in the "Admission" table. Focusing on these records ensures that the study targets the most urgent care scenarios. Patients with diabetes are then isolated using a more robust method involving SUM(CASE statement combined with UPPER() and TRIM() functions. This approach enhances the query's resilience by accounting for case insensitivity and removing any extraneous spaces, thus ensuring accuracy in identifying diabetic patients regardless of future variations in data entry practices. (Khalil)

To quantify the impact, the filtered data are aggregated, counting the instances of diabetic patients admitted through emergency services. This metric is vital for assessing the demand for critical care resources and guiding policy enhancements in emergency healthcare services.

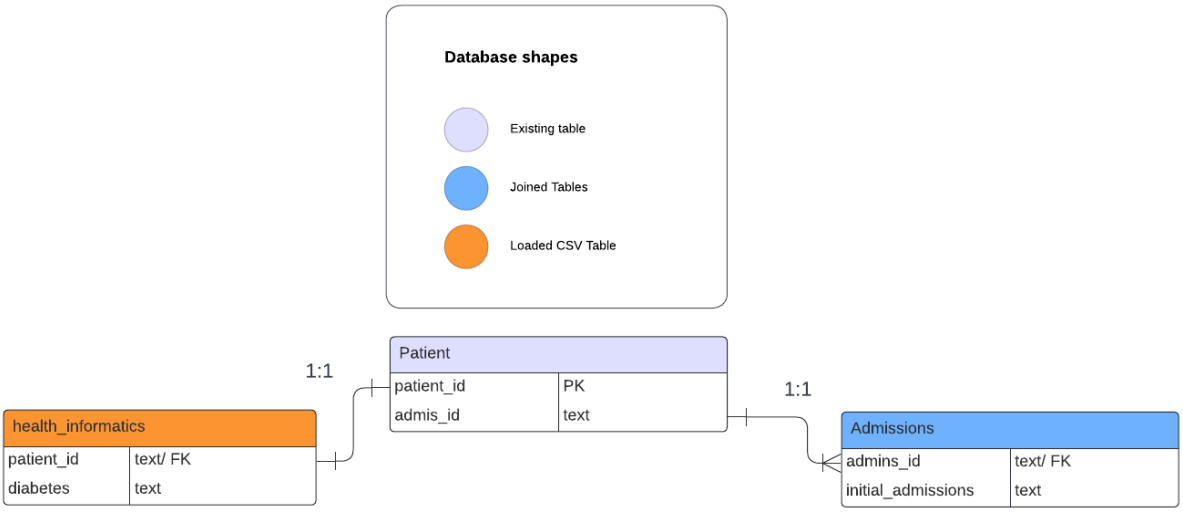
This methodological approach leverages sophisticated data handling techniques to provide a precise count of emergency admissions for diabetic patients, thereby supporting enhanced healthcare planning and policy formulation. The integration of comprehensive data sources with advanced SQL querying ensures both the accuracy and relevancy of the analysis in a clinical context.

2. Relevant Data Fields

Table	Column	Data Type	Key
Patient	patient_id	Text	Primary
Patient	admis_id	Text	
Health_informatics	patient_id	Text	Foreign
Health_informatics	diabetes	Text	
Admission	admins_id	Text	Foreign
Admission	inital_admissions	Text	

Each field has been selected to ensure accurate matching and filtering of patient records based on diabetes status and the type of admission, particularly focusing on emergency services. This setup facilitates the combination of data from both the original

database and the externally provided health informatics CSV file, effectively answering the research question through a structured query.



Section B:

There where many other columns in each table but these columns were not relevant for questions answered in part A. The diagram includes all columns used in the analysis and there relationship to each other.

1.

Relationship Between the Patient Table and Health_informatics Table

The Patient table is directly linked to the Health_informatics table via a one-to-one (1:1) relationship. This relationship is defined by the patient_id, which is used as a primary key in the Patient table and as a unique identifier in the Health_informatics table. Each record in the Patient table corresponds to a single, unique record in the Health_informatics table, which contains detailed health information about each patient.

Relationship Between the Patient Table and Admission Table

The current structure of the relationship between the "Patient" table and the "Admission" table in our database is characterized as a one-to-one (1:1) relationship. Within the "Patient" table, each individual is uniquely identified by a patient_id and linked to a singular admission_type value, designated as 1, 2, or 3. These values correspond to distinct types of admissions—namely, emergency, elective, or observation—each of which is further detailed in the "Admission" table. The "Admission" table functions as a lookup resource, clarifying the implications of each admission_type value associated with a patient. (Khan)
.

Under typical circumstances, the "Admission" table would facilitate a one-to-many (1:M) relationship, reflecting multiple admissions per patient over time. However, due to the "Patient" table's configuration, which records only a single admission_type for each individual and lacks a datetime column to track multiple events, the relationship is effectively constrained to one-to-one.

2.

— Create health_informatics Table.

```
DROP TABLE public.health_informatics
```

— Check data got deleted from the Table.

```
SELECT count(*)
FROM health_informatics
```

— Create health_informatics Table.

```
CREATE TABLE public.health_informatics
(
    patient_id text NOT NULL,
    services text NOT NULL,
    overweight text NOT NULL,
    arthritis text NOT NULL,
    diabetes text NOT NULL,
    hyperlipidemia text NOT NULL,
    backpain text NOT NULL,
    anxiety text NOT NULL,
    allergic_rhinitis text NOT NULL,
    reflux_esophagitis text NOT NULL,
    asthma text NOT NULL,
    CONSTRAINT patient_id_key PRIMARY KEY (patient_id)
)

TABLESPACE pg_default;

ALTER TABLE public.services
    OWNER to postgres;
```

— Upload Date from CSV

```
COPY health_informatics
FROM 'C:\LabFiles\Medical\mservices.csv'
DELIMITER ','
CSV HEADER;
```

--Check data got uploaded to created table.

```
SELECT count(*)
FROM health_informatics
```

Section C:

How many patients with diabetes are admitted to emergency services?

```
SELECT
        initial_admission,
        sum(CASE WHEN UPPER(TRIM(diabetes)) = 'YES' then 1 else 0 end)
has_diabetes
FROM patient p
LEFT JOIN health_informatics h on h.patient_id = p.patient_id
LEFT JOIN admission a on a.admins_id = p.admis_id
group by initial_admission
```

** File was uploaded with assessment labeled as result.csv

Section D:

Given the one-to-one relationship between the `Patient` table and the `Health_informatics` table, it is prudent to synchronize the refresh rate of the `Health_informatics` add-on file with the rate at which patients are admitted. Therefore, the optimal update frequency for the `Health_informatics` table is on a daily basis.

Relevance to Business Activities

Updating the `Health_informatics` table daily ensures that the information reflects the most current admissions, which is especially important for tracking the admissions of patients with diabetes through emergency services. Since this table captures the health status of patients as they are admitted, it provides a timely and accurate representation of their current health metrics, including their diabetic status.

For a healthcare provider monitoring the effectiveness of interventions aimed at reducing diabetic emergencies, having access to up-to-date health informatics is crucial. It allows for the immediate assessment of whether the frequency of diabetic emergencies is increasing or decreasing, which in turn informs the tactical decisions around patient education and preventive care programs.

A daily refresh ensures that any changes in a patient's health status are promptly reflected in the database, allowing healthcare providers to react swiftly to trends and potentially modify patient care plans to better manage their diabetes. This approach is aligned with value-based care objectives, which focus on improving patient outcomes and reducing unnecessary healthcare expenditures. By maintaining an up-to-date database, providers can more accurately measure the impact of their care strategies and adjust them to optimize patient health outcomes.

References:

Khalil, M. (n.d.). *Case when with aggregate functions: SQL*. campus.datacamp.com. https://campus.datacamp.com/courses/data-manipulation-in-sql/well-take-the-case?ex=8

Khan, M. F. (n.d.). Defining relationships: SQL. campus.datacamp.com. https://campus.datacamp.com/courses/joining-data-in-sql/introducing-inner-joins?ex=5