D205: Data Acquisition Performance Assessment

1. **Research Question**

Hospitals are trying to prevent patients from being readmitted because “readmission rates are significantly influencing a hospital's quality scores and performance” (Fallon, 2023). Patient’s medical conditions have been provided to reflect whether certain conditions influence readmission more than other conditions. However, this raises a concern on if readmission is more likely in a specific age group rather than a medical condition. Therefore, I would like to focus on senior citizens and their survey responses to figure out if there is something the hospital could improve on to decrease readmission in this specific age group. My research question is the following: ‘Is there a factor from the survey responses that is more important than the other factors that could decrease senior citizens patients from readmission?’

**A1. Identifying Data**

The research question chosen is in regard to the correlation between a specific age group,

being senior citizens, patient readmission, and their responses to the survey they received. On

that note, the data needed to answer the research question is the following:

Table Name: patient

Column Names: patient\_id character varying (50), age integer , readmis text

Primary Key: patient\_id character varying (50)

CSV: survey\_responses

Table Name: survey\_responses

Column Names: patient\_id text, timely\_admission integer, timely\_treatment integer,

timely\_visits integer, reliability integer, options integer, treatment\_hours integer, courteous

integer, active\_listening integer

Primary Key: patient\_id character varying (50)

Foreign Key: patient\_id character varying (50)

1. **A screenshot of a computer

   Description automatically generatedEntity Relationship Diagram (ERD)**

There is a 1:multiple relationship between patients and survey responses since there are multiple questions each patient had to answer.

**B1.** **Code for the ERD**

CREATE TABLE public.survey\_responses

(

patient\_id character varying(50) NOT NULL,

timely\_admission integer,

timely\_treatment integer,

timely\_visits integer,

reliability integer,

options integer,

treatment\_hours integer,

courteous integer,

active\_listening integer,

PRIMARY KEY (patient\_id)

);

ALTER TABLE public.survey\_responses

ADD FOREIGN KEY (patient\_id)

REFERENCES public.patient (patient\_id);

**B2. Loading CSV Data**

After creating a table for the survey responses from the patients, the data loaded from the

survey responses add-on file was the following:

"\\copy public.survey\_responses (patient\_id, timely\_admission, timely\_treatment, timely\_visits, reliability, options, treatment\_hours, courteous, active\_listening) FROM 'C:/LabFiles/Medical/msurvey.csv' DELIMITER ',' CSV HEADER QUOTE '\"' ESCAPE '''';""

1. **SQL Query**

The following code was queried in pgadmin to join the survey\_responses.csv file and the patient information:

SELECT p.patient\_id, p.age, p.readmis

FROM patient AS p

LEFT JOIN public.“survey\_responses” AS s

ON p.patient\_id = s.patient\_id;

The following code was queried to answer my research question, “Is there a factor from the survey responses more important than the other factors that could decrease senior citizen patients from readmission?” The constraints are patients 65 years old and older as well as if the patient has been readmitted. I found the sum of each survey response factor to compare the results. The lowest sum will be the most important survey response factor since 1 is the most important and 8 is least important.

SELECT SUM(timely\_admission) AS timely\_admission,

      SUM(timely\_treatment) AS timely\_treatment,

      SUM(timely\_visits) AS timely\_visits,

      SUM(reliability) AS reliability,

      SUM(options) AS options,

      SUM(treatment\_hours) AS treatment\_hours,

      SUM(courteous) AS courteous,

      SUM(active\_listening) AS active\_listening

      FROM patient AS p

      LEFT JOIN public."survey\_responses" AS s

      ON p.patient\_id = s.patient\_id

      WHERE age >= 65 AND readmis = 'Yes';

**C1. CSV Files**

I have attached the csv file with the results for my research question to the submission.

**D.  Add-On File**

The medical hospital chain should be updating and refreshing the add-on file daily. This

includes any time new data from the external sources is modified, whether that be changed to

different information or additional information.

**D1. Explanation of Time Period**

When the data is updated/modified, it has to be refreshed daily. That way, there will be

enough data accumulated from the patients that day which could possibly change the

results of the research question. This could in turn figure out which factor the hospital could

improve on to reduce their readmission problem.

**E.  Panopto Video of Code**

"\\copy public.survey\_responses (patient\_id, timely\_admission, timely\_treatment, timely\_visits, reliability, options, treatment\_hours, courteous, active\_listening) FROM 'C:/LabFiles/Medical/msurvey.csv' DELIMITER ',' CSV HEADER QUOTE '\"' ESCAPE '''';""

**E1. Panopto Video of Programs**

**F.  Web Sources**

Youtube. (2019, May 27). *Intro to postgresql databases with pgadmin for beginners – full course.* Youtube. https://www.youtube.com/watch?v=Dd2ej-QKrWY

**G.  Sources**

Fallon, C. (2023, June 27). 4 reasons why a hospital readmission rate matters. Blog. https://blog.cureatr.com/4-reasons-why-a-hospital-readmission-rate-matters#:~:text=There%20is%20a%20correlation%20between,for%20hospitals%20and%20their%20leadership.

**H.  Professional Communication**

Demonstrate professional communication in the content and presentation of your submission.