Aryaman Babber

CS 301

Sql Project Report

**QUESTION 2**

Question 1

Calendar

Description automatically generated

I first added an int variable called “age” to the census table. Then, I updated the variable “age” to equal 2020 minus the substring of the id number starting at index 6, of length 4.

Question 2

Calendar

Description automatically generated

I added a float variable “bmi” to census.

I created a temp table where I left joined census and patient\_info, with all the attributes from census, and the attributes height\_cm and weight\_kg from patient\_info.

Then, I updated bmi in the temp table using the equation given in the guidelines.

I dropped the height\_cm and weight\_kg columns in the temp table, and dropped the original census table. I then created a new census table to replace the old one, which was equal to the temp table.

Question 3

Table

Description automatically generated

A picture containing calendar

Description automatically generated

We create a temp\_census table that left joins all of census to the needed attributes from patients\_physical\_exams.

We create temp\_2 table that inner joins all of temp\_census to the max(exam\_date) of each id\_Number in temp\_census. Now, temp\_2 has chosen the most recent exam dates for each id\_Number (each person).

After the last step, there are 24 duplicates remaining, from people who took a physical exam twice on the same day. We create temp\_3 table by inner joining temp\_2 and max(metabolism). This gets rid of the 24 duplicates, and chooses the exam where the patient had a higher metabolism rate.

We create and set a fitness\_score variable to temp\_3 using the equation in the guidelines. We drop temp\_census. We create a new temp\_census where we left join census with temp\_3. Now, temp\_census contains all attributes from census, with a fitness score of everyone who has taken a physical exam.

We update fitness\_score where fitness\_score equals Null to the average score from temp\_3. Now, all of temp\_census has a fitness score, either from their physical exam, or the average fitness score. We see that in tables 1 and 2 shown above.

Question 4

A picture containing table

Description automatically generated

We create a temp\_1 table that left joins the census fitness score to all of audiences.

We create a temp\_2 table that left joins all of temp\_1 and the zovid12test test\_result.

We drop audiences, and create a new audiences set to temp\_2.

**QUESTION 3**

Question 1

Text

Description automatically generated

Question 2

Table

Description automatically generated

Integration Explanation

First, it is important to note that I used Jupyter Notebook (Python 3 notebook) for this project.

Using pip install, I downloaded the sqlalchemy library, from which I could import sqlalchemy.engine.



I then pip installed my\_sql\_ext for Python, which allows for sql magic in a Jupyter notebook. Using this, I loaded the sql extension into the notebook. I then used a sql query to access the mysql database.

Text

Description automatically generated

This connects my sql database to python for the rest of the notebook, so I can use sql queries later on.