

INFO20003 Semester 2, 2019

Assignment 2: SQL

Due: 6:00pm Friday 20th of September

Submission: Via LMS <https://lms.unimelb.edu.au>

Wellbeing Application

You are the SQL programmer for a mental wellbeing app, which allows users to do two types of things. The first is to participate in a Twitter-style social network. The second is to do online units called Steps, which are interactive therapy modules designed to exercise and develop a range of psychological skills.

How the app works

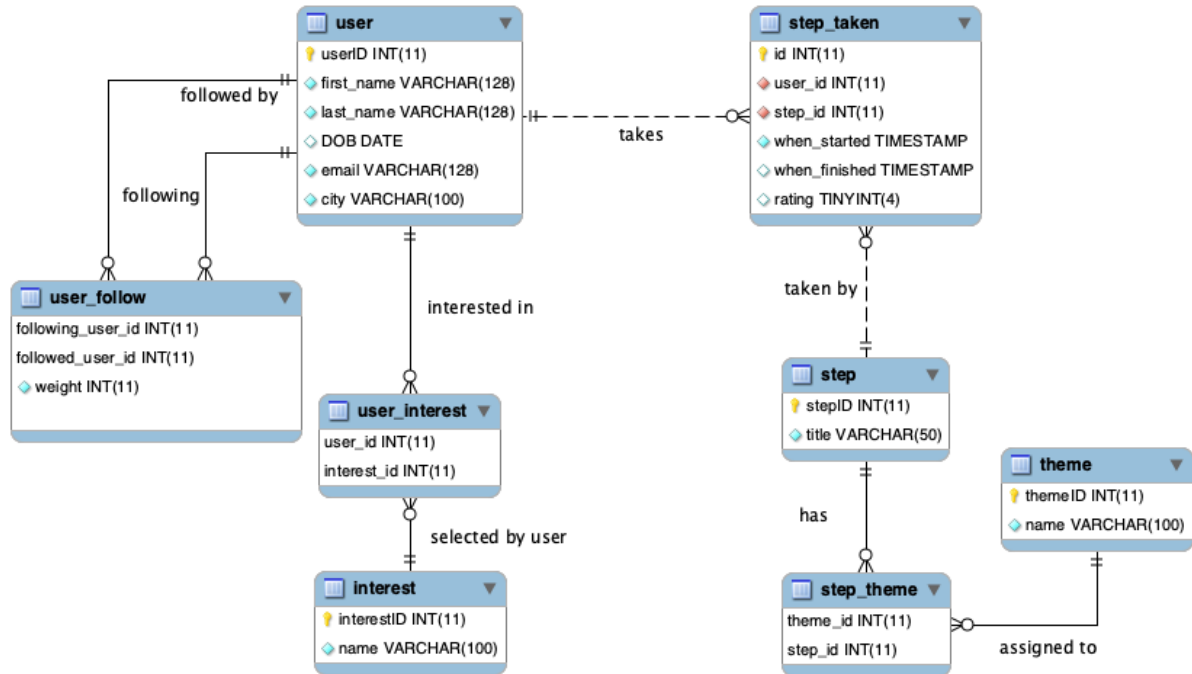
When users register into the system, a record is created in the 'user' table. Upon registering, a user can follow other users to see their posts and activity. If user x (following) starts following user y (followed), an appropriate entry is created in the 'user_follow' table (x following y does not necessarily mean that y follows x). The weight attribute of this table holds an integer that reflects the extent to which x follows y ; for example, this number would increase the more x likes or comments on posts created by y . To help connect users, each user can also select, from a list stored in the 'interest' table, interests that they have (e.g. movies, music, books). For each interest that a user selects, a record is created in the associative table 'user_interest'.

On the therapy side of this app, users can do online modules called Steps. Different Steps take different amounts of time to complete and users can do a Step in one sitting or return to a Step and complete it later on. Sometimes they might not finish a Step and will instead move on to something else. Also, they can do any given Step as many times as they would like and each time, they can rate how it made them feel (1 being the lowest and 5 being the highest).

As soon as user x makes a start on doing Step z , an entry is created in the 'step_taken' table to record the fact that x started doing z . The timestamp of this start is also automatically recorded in the 'when_started' column. If the user successfully completes the Step, the datetime of completion is recorded in the 'when_finished' column.

Steps are also categorised under themes, which are used to indicate what types of things a Step might help with. For example, a Step that helps with mindfulness would be categorised under the 'Mindfulness' theme. A given step can be categorised under more than one themes.

Data Model



Setup Script

A dataset is provided against which you can test your solutions to the assignment. To set up the dataset, download the file *wellbeingapp.sql* from the Assessments folder on LMS and run it in MySQL Workbench. This script creates the database tables and populates them with data. The script is designed to run against your account on the Engineering IT server (*info20003db.eng.unimelb.edu.au*). If you want to install the schema on your own MySQL Server installation, uncomment the BYOD section at the beginning of the script.

The SQL Tasks

In this section are listed 10 questions for you to answer. Write one (single) SQL statement per question. Subqueries and nesting are allowed within a single SQL statement – however, you will be penalized for writing overly complicated SQL statements. DO NOT USE VIEWS to answer questions.

1. Find all the ongoing/unfinished steps. Display the title of these steps and full names of the users who are taking these steps. **(1 mark)**
2. List the themes and the number of the steps associated with these themes. Display the theme name and number of associated steps sorted in descending order. **(1 mark)**
3. Which step is the least popular based on the average rating given by users? Display the title and ID of the step and its average rating (formatted to 2 decimal places). Only include those steps which are rated by at least one user. **(1 mark)**
4. Find the steps that are taken the greatest number of times. Display the ID, title and count of the times the step has been taken. In case of ties, display all the steps with the same number of times taken. **(2 marks)**
5. Who is the most followed user between age of 15 and 18? Display the age (as an integer), first name, and last name of such user along with the number of followers. **(2 marks)**
Hint: The function `TIMESTAMPDIFF` can be used to subtract two timestamps. The function `CURDATE` returns current date.
6. Find all steps that are never taken or are taken exactly once? Display the id and title of these steps along with the indication how many times the step has been taken (0 or 1). **(2 marks)**
7. Find users who started taking step 'Doing and being' after they had started the step 'Panic' but have never completed 'Panic'. Display the user ID, first name and last name. **(2 marks)**
8. What finished steps were completed both by a user with first name "Alice" and a user with first name "Bob"? Display the ID and title of such steps along with the number of times each user has completed these steps. **(3 marks)**
9. Find the top two users with the highest number of interests. For those two users, find out the common steps taken by both of them. Display the titles of the common steps they have taken and the number of times those steps are taken by each user. **(3 marks)**
10. For each user taking a step, calculate how many other users have taken the same step. We are only interested in the cases where the step is performed by at least 5 other users. Display the user ID, number of other users (at least 5 other users) who are taking the same step and the title of the taken step. **(3 marks)**

Submission Instructions

Submit a single PDF showing your answers to all questions to the Assessment page on LMS by 6pm on the due date of **Friday 20th of September**. Name your file 987654.pdf, where 987654 corresponds to YOUR student id. Other formatting requirements are listed below.

For each question, present an answer in the following format:

- Show the question number and question in black text.
- Show your answer (the SQL statement) in **blue text (not a screenshot)**.
- Show a screenshot from Workbench containing the output of the query.
 - If the query returns more than 10 rows, take a screenshot of only the first 10 rows.
- Show how many rows were actually returned in **red text**.
- Show each query on a **separate page**.

Example:

QXX. List the names of the themes.

```
SELECT name
```

```
FROM theme;
```

name	
▶ Social Anxiety	
Your Strengths	
Meditation	
Depression	
Anxiety	
Self Compassion	
Mindfulness	

7 rows returned

Requesting a Submission Deadline Extension

If you need an extension due to a valid (medical) reason, you will need to provide evidence to support your request by *9pm, Thursday 19th of September*. Medical certificates need to be at least two days in length.

To request an extension:

1. Email Farah Khan (khanf1@unimelb.edu.au) from your university email address, supplying your student ID, the extension request and supporting evidence.
2. If your submission deadline extension is granted you will receive an email reply granting the new submission date. Do not lose this email! Replies may take up to 12 hours, so please be patient.

Reminder: INFO20003 Hurdle Requirements

To pass INFO20003 you must pass two hurdles:

- Hurdle 1: Obtain at least 50% (15/30) or higher for the three assignments (each worth 10%)
- Hurdle 2: Obtain a grade of 50% (35/70) or higher for the End of Semester Exam

Therefore, it is our recommendation to students that you attempt every assignment and every question in the exam.

GOOD LUCK!