

INFO 6210

Data Management and Database Design

Practice Exam One SQL

Student Name: _____
Professor: Nik Bear Brown

Rules:

1. NO COMPUTER, NO PHONE, NO DISCUSSION or SHARING.
2. Ask if you don't understand a question.
3. Time allowed. Last 90 minutes of class
5. Bring pen/pencil. The exam will be written on paper.

All of the questions in the exam refer to the Twitter database shown below

```
CREATE TABLE IF NOT EXISTS "tweets" (  
    "tweet_id"      varchar(25) NOT NULL,  
    "user_id"       varchar(25) NOT NULL,  
    "tweet_content" varchar(300) NOT NULL,  
    "tweet_datetime" datetime NOT NULL DEFAULT CURRENT_TIMESTAMP,  
    "favorites"     int,  
    "retweets"      int,  
    "urls"          varchar(100),  
    PRIMARY KEY("tweet_id")  
);
```

```
CREATE TABLE IF NOT EXISTS "users" (  
    "user_id"      varchar(25) NOT NULL,  
    "user_name"    varchar(50) NOT NULL,,  
    "followers"    int,  
    "following"    int,  
    "tweet_count"  int,  
    PRIMARY KEY("user_id")  
);
```

```
CREATE TABLE IF NOT EXISTS "tags" (  
    "tag_id"        integer NOT NULL PRIMARY KEY AUTOINCREMENT,  
    "tag"           varchar(50) UNIQUE  
);
```

```
CREATE TABLE IF NOT EXISTS "tweet_tags" (  
    "tweet_id"      varchar(25),  
    "tag_id"        varchar(25)  
);
```

Q1 (5 Points)

Show the user_id, user_name. and followers of the top 5 users by number of followers.

Q2 (5 Points)

Get the top 7 days with the highest total number of retweet..

Q3 (5 Points)

Create a column that is calculated from other columns in the tweets table.

Q4 (5 Points)

Count all of the null values in a nullable field.

Q5 (5 Points)

Computationally what is the most expensive operation in the relational data model?

Q6 (5 Points)

Write a function to get the numbers of retweets for a given user_id.

Q7 (5 Points)

Why not put all the data in one big table and avoid all of these joins?

Q8 (5 Points)

Why create views?

Q9 (5 Points)

Show the tweet_id, user_name, and tweet length of the 5 longest tweets.

Q10 (5 Points)

Create a view called 'longest_tweets' using the SQL in Q9

Q11 (5 Points)

Why create temporary tables?

Q12 (5 Points)

Write SQL to insert a new user into the user table.

Q13 (5 Points)

Create a temporary table called 'longest_tweets_tmp' using the SQL in Q9.

Q14 (5 Points)

Update the user you added in Q12 to have 55 following.

Q15 (5 Points)

Update all users with null or negative followers to have 0 followers.

Q16 (5 Points)

Delete a user with the user_id of '@Spammy' and then check that the user is removed.

Q17 (5 Points)

Create a procedure called tag_match that shows the tags in the tags table that start with a given text string.

Q18 (5 Points)

Create a index on a non-key attribute.

Q19 (5 Points)

Create a table called 'user_log' with the user id and a log_date field that represents the current date and time. The primary key is the two fields in the table.

The create a trigger that will use the user_log table to record the user_id and current date and time when a new user is inserted into the user table.

Q20 (5 Points)

Create a transaction that runs the SQL in Q16. That is, delete a user with the user_id of '@Spammy' but only finalize the SQL if only one user is removed from the users table.

Update the SQL check to make sure that it counts how many users were removed if it isn't already doing that. Only complete the transaction if only one user is removed

Assume you are doing this from the command line but can write a stored procedure to do this if you wish.