Homework 6: More SQL (Hands-on)   (100 points)

Due Date: Friday, Nov 20 (4:00 PM)

Submission

All HW assignments should contain both your student ID and your name and must be submitted online via the HW6 dropbox on Gradescope. See the table below for the HW 6 submission opportunities. Note that after 4 PM on November the 21st no further HW 6 submissions will be accepted. (We will be releasing the solution at that time.) Please strive to get all your work in on time! If possible, try to save the one dropped assignment for the end of the term when you are most likely to want/need it.

Date / Time	Grade Implications
Friday, Nov 20 (4:00 PM)	Full credit will be available
Saturday, Nov 21 (11:00 PM)	10 points will be deducted

More Structured Query Language (SQL) [30 pts]

We have updated the datasets used in HW5. Please **reload the new data** using the script from here (again, you must use your UCI email to download this). **Refer to the post here if you are having issues loading the new data.** You are going to use MySQL for all of the queries in the assignment and turn in the queries and results according to the provided template.

Write the following DML and DDL statements in SQL against the new *CheckedTweets* test database. Show the result of each query where requested to do so. Please note that you will not get points for providing the result of a query on this assignment if your SQL query is syntactically incorrect (i.e., if it doesn't execute). Since you have a "live" system at your disposal, this should not be an issue – you can easily run all of your queries that way. **Also, make sure that the results of your queries do not contain duplicate records, as you will lose points if they do.** Note that some of the problems' solutions will have side effects on the database; you can use other SQL statements to undo those effects while debugging, and in the worst case you can always reload the data if you inadvertently mess it up while getting everything working. When you run your answers to turn them in, **start with a clean load** and run them all in sequence.

NOTE: For questions requiring you to write DDL statements (e.g., views, triggers, stored procedures, alter table), you must write the required DDL statements **by hand**. You will **not** get the points for a problem if you let the GUI tool generate the DDL for you!

ALSO NOTE: Read the above **NOTE** one more time. We're dead serious about this!

Appendix

1) To minimize any copy-paste issues from the template to MySQL Workbench, we recommend copying any queries found in the template from the provided script here instead.

2) In addition to the load script specifying the structure and data you will be working with, we have also defined a stored procedure for you (which you will need for question 7) at the tail end of the load script itself. For your reading enjoyment, the provided procedure is also shown here (as another example of a stored procedure for you to see):

```sql
DELIMITER //
CREATE PROCEDURE UpdateHashtags( raw_tweet_id VARCHAR(20) )
BEGIN
    INSERT INTO Hashtags
    SELECT hashtags.*, T.tweet_id
    FROM RawTweet T,
         JSON_TABLE(
             T.content,
             '$.extended_tweet.entities.hashtags[*]'
             COLUMNS(hashtag_text VARCHAR(50) PATH '$.text')
         ) hashtags
    WHERE T.tweet_id = raw_tweet_id;
END; //
DELIMITER ;
```