Database Systems, CSCI 4380-01
Homework # 8
Due Monday October 28, 2019 at 11:59:59 PM

Homework Statement. This homework is worth 3% of your total grade. If you choose to skip it, Midterm #2 will be worth 3% more. This homework is on writing insert/update/delete statements.

Your homework will use the same dataset as Homework 7. Since you are changing data, you will use your own database (`db_username`), created already for you and populated with this data.

There are only a few rules:

- Do not create tables or any other objects in any other database, other than your own.
- Do not change the already given tables as it will be costly for me to recreate them. Do all your work by creating new database objects as described below.
- Only create the tables described and nothing else. You can use more than one insert/update/delete statement for each expression, but do not use additional tables to simplify the expressions.
- Use only the SQL standard for `UPDATE` statement. Do not use the Postgresql syntax, i.e. `UPDATE SET WHERE...`. This is not standard and is kind of confusing.

Homework Description

Write the following expressions in SQL. Note that each expression must be evaluated in the given order. Use the given template as a starting point.

First, create a copy of the listings/reviews/calendar relations, and remove your ability to change the base tables to avoid any mistaken changes:

```
revoke insert,update,delete on calendar from <your_username>;
revoke insert,update,delete on hotel_regions from <your_username>;
revoke insert,update,delete on hotels from <your_username>;
revoke insert,update,delete on listings from <your_username>;
revoke insert,update,delete on rental_prices from <your_username>;
revoke insert,update,delete on reviews from <your_username>;
revoke insert,update,delete on ta_reviews from <your_username>;

create table reviews2 as select * from reviews;
create table listings2 as select * from listings;
create table calendar2 as select * from calendar;
create table hotel_regions2 as select * from hotel_regions;
create table hotels2 as select * from hotels;
```

Expression 1 Find all listings that has less than 5 reviews (`number_of_reviews` and the last review was more than 5 years ago (according to `last_review` attribute). Delete all tuples in reviews2/calendar2/listings2 for these listings.

Expression 2 We are going to do some data cleaning for the `hotels2` table.

If the zip value is equal to the empty string, set it to NULL.

For the remaining entries, if the ZIP value is longer than 5, then set it to the first 5 characters of the current ZIP value. Note that some entries contain a leading `'NY '` which you should also trim.

Expression 3 Add a new column called latest review count as shown below:

```
alter table listings2 add latest_number_of_reviews integer;
```

Update the listings table by setting the latest review count to be the number of reviews of that listing in 2019. Remember, if there are no reviews, then the count should be zero.

Note: If you try to set every listing count with a single query, it will take too long. Instead, set the number to zero for all tuples first. Then, only set the count to a non-zero value for the listings that has at least one review in 2019.

Expression 4 The following zip codes have a single region in `hotel_regions` but more than 10 hotels: `10021, 10022, 10024, 10025`. We assume the region is fixed for all these zip codes.

Insert into `hotel_regions2` all region info for all hotels `H` such that:

- `H` is in one of the zipcodes `10021, 10022, 10024, 10025` (you can hardcode these zip code values),
- the id for `H` is not currently in `hotel_regions2`, and
- the region for `H` is the same as the region for any of the hotels in `hotel_regions2` with the same zip code.

For example, hotel `93466` is not currently in `hotel_regions2` and has zipcode `10022`. Hotels in `hotel_regions2` with zipcode `10022` has the region/county `Midtown, Manhattan`. So, the tuple `(93466,Midtown, Manhattan)` must be one of the new tuples we insert.

SUBMISSION INSTRUCTIONS. You will use SUBMITTY for this homework.

Use the template SQL file given to you that shows you where to insert your answers to the given expressions. After filling in the code and putting in your username where it says `<your_username>`, rename the file `username_hw8ans.sql` and submit it on SUBMITTY.