Database Systems, CSCI 4380-01
Homework # 5
Due Thursday March 24, 2011 at 2 pm

You are given attached schema which is the same one used for Homework #4. Recall that, this database allows people to enter information about people, places and businesses. It is fairly free form. However, we store each version of information for people using some sort of versioning. Each table locations, businesses and people have an additional version table (name ends with V) which stores the past versions of each location, business and people as well as the most current version. If information is deleted from the main table, it is also deleted from the version table.

We also allow users to state relationships between people, places and businesses. Each relationship can only be stated once, so we store who first established this relationship. Finally, people are able to review these records, post comments on reviews and comments that follow up other comments in a thread. For each comment, we store the parent review it links to.

Users can also friend others, vote on comments up or down (but not vote on reviews, which is a limitation of the system).

Write the following using SQL. Each query is worth 15 points. **Note:** Choose any six to answer. If you answer all 7, you will get a bonus 15 points.

(1) Find all locations entered by a user living in the same zipcode as that location (use v0 for that location).

(2) Return all tuples in 'Locations' as the last version of tuples in 'LocationsV'. In other words, if the location is in LocationsV, then it should have 1 + the highest version stored LocationsV. If the location is not in LocationsV, then it should have version 0.

(3) Find the most contraversial people, i.e. people with the highest number of versions. Return their id, first and lastname.

(4) Find businesses that are at the same location and have the same type. (Use the latest version of business and location for this query). Return the businesses and the location and type they share.

(5) Find businesses that have a relation to the same person both directly (in BPRel relation) and indirectly (through LBRel and LPRel relations). Return the name of the business and the person.

(6) Find all mutual friends (i.e. they have requested each other) who also share interests: either they updated the same business or the same person. (Use only the businesssesV and peopleV relations).

(7) For each user, find the total number of reviews, the average value of their reviews, the number of comments they received (through their reviews), the total value and standard deviation of votes they received to their comments. Return the user id, name and the asked values.

1 Deliverables

Turn in a single text file (.sql) containing all your queries. It must be possible to execute the whole file using the

`\i filename`

command in psql. To achieve this, make sure all queries execute and end with ;. If you have a query that does not run, you will loose all points for that query. It is better for you to turn in a query that is not fully correct than one that does not run. Also, note that you can write commands in a .sql file by preceding it with `--` . So, comment each query by preceding it with a line that describes the query you are answering. You should also use the psql command

`\echo 'text'`

to identify your name at the top and then the id of each query before the SQL for it. This will allow your TA to quickly run and test your homework. A template for answers is provided with this homework. You must use this template and fill in your personal inforamtion and the SQL queries.

To test your queries, connect to the postgresql server at CS. First ssh to remote.cs.rpi.edu.

`ssh remote.cs.rpi.edu -l username`

using your CS username and start postgresql using the username and password mailed to you and the database csc4380.

`psql -h csc4380.cs.rpi.edu csc4380_wud6 -U csc4380_username`

again using your CS username. You can get help on how to use postgresql using the online documentation:

`http://www.postgresql.org/docs/8.2/interactive/index.html`

If you are using the CS systems for your queries, you can run the file containing your answers in the Unix prompt with the following comment:

`psql -h csc4380.cs.rpi.edu csc4380_wud6 -U csc4380_username -f answer_file > output_file`

This will run the queries in the given `answer_file`, and store the results in a new file called `output_file`.