100 Points

Due September 24th, by 5 pm, electronically using SPARK.

In this assignment, you will write SQL queries to be executed over data from facebook. This data will obtained from your own facebook profile! The schemas of the tables are described below:

- **Person** (pid:bigint, name:varchar(50), first_name:varchar(30), last_name:varchar(30), birth-day:varchar(30), hometown:varchar(50), current_location:varchar(50), gender:varchar(10), relationship_status:varchar(20))

- **TaggedPhotos** (tagged_person:bigint, tagged_by:bigint references Person.pid, photoid:bigint references Photos.photoid, created_time:timestamp, link:varchar(256))

- **PhotoAlbums** (owner:bigint, albumid:bigint, albumname:varchar(100), location:varchar(256), count:integer)

- **Photos** (photoid:bigint, albumid:bigint, created_time:timestamp, link:varchar(256))

- **Likes** (pid:bigint, name:varchar(50), category:varchar(50), pageid:bigint)

TaggedPhotos.tagged_person is a foreign key referencing Person.pid.
PhotoAlbums.owner is a foreign key referencing Person.pid.
Photos.albumid is a foreign key referencing PhotoAlbums.albumid.
Likes.pid is a foreign key referencing Person.pid.

1. **Obtaining data and populating tables**

In the first part of the assignment, you will download your data using our facebook application. Then, you will load this data into your own database, using the Postgres database server, which is running on edlab machines. Instructions for obtaining the data, loading the data, links to Postgres documentation, and other resources are available on the course website under ‘System Support’ tab.

2. **Practical SQL exercises**

Now, write SQL expressions for each of the following queries and execute them. Please submit for each query below (i) a valid SQL query, and (ii) the output of the query on the sample data – not your personal facebook data.

(a) Find the distinct names of the persons who are tagged in at least one photo.

(b) For each tagged photo, print its photoid, albumname and created_time. Limit the result set to 10.
(c) Find the distinct names of the persons who share at least one page liked by you. (Hint: You will need to use conditions in the 'WHERE' clause which depend on your entry in the database. You can find your entry as the first row in the 'person.txt' file downloaded using the facebook application)

(d) For each category in Likes, print the 'category' along with the number of distinct pages for that category.

(e) For each person, print the name of the person along with the total number of photos he/she has uploaded. It is fine if your query does not print names of the persons who haven't uploaded any photos.

(f) Find the names of the persons who like a page in category 'Interest' and a page in category 'Movie'

(g) Find the names of the persons who like a page in category 'Interest' or a page in category 'Movie'

(h) Good Dating Candidates: List pairs of your friends where both have relationship status 'Single' and they share at least one liked category in common. If you wish, you can put gender constraints on friends being matched to improve the result quality.