

Assignment Three: Nested Subqueries in SQL

Assignment three is to use sql on agora using the Postgresql database system and the psql command interpreter to develop a series of queries involving nested subqueries. In particular, you will submit a tar file that contains the following files

- a psql batch file you write that has the DDL for dropping all the tables and has at the top as a comment your names and date. The two characters -- (with no space in between) start a comment in a psql batch file. Remember that the order in which you drop tables has to be consistent with your foreign key constraints.
- a psql batch file you write that has the DDL to create all the tables and has at the top as a comment your names and date. Remember that the order in which you create tables has to be consistent with your foreign key constraints.
- a psql batch file you write that uses the \copy psql command to populate all the tables and has at the top as a comment your names and date. Suppose you want to populate the file named blue from a file in the data subdirectory of the assg3 directory, then the \copy command in your batch file is
`\copy blue from 'data/blue_data.txt';`
 assuming that you are running that batch file while you are in the assg3 directory. Remember that the order in which you populate tables has to be consistent with your foreign key constraints.
- a psql batch file you write that has the DML select statements and \echo statements for your answers to the problems. I have provided a template for your file that has the \echo statements including \echo statements for the expected results with the provided example database instance. The template is called assg3_tester.sql.
- a readme file that includes your author names, the data of submission, a description of what is the assignment files and that you have done, and if any of your answers you do not think is completely correct an identification of the exercise number and a statement about what you think is not working correctly.

Details

- Just below this handout on the course website is a copy of the assg3_tester.sql file that your DML batch file has to be an extension of. In particular, for each query where the assg3_tester.sql file has a
`\echo 'replace this line (including \echo) with your query'`
 you should replace that line with your query.
- The database schema to use for the DDL and to answer your queries is below. The domain for each attribute is shown, though the actual SQL name for the domain might not be exactly that word. For example, the domain string, you would replace by varchar(20) or something like that. The attributes that form the primary key are underlined.
 - sailors(sid:integer, sname:string, rating:integer, age:real)
 - boats(bid:integer, bname:string, color:string)

- `reserves(sid:integer, bid:integer, day:date)`
- Each of your queries to the problems must include in a fundamental way at least one nested subquery.
- At least one nested subquery must use the IN connective.
- At least one nested subquery must use the NOT IN connective.
- At least one nested subquery must use the EXISTS connective.
- At least one nested subquery must use the ALL or ANY connectives. SOME is a synonym for ANY, so you can use SOME instead of ANY.
- At least two correlated subqueries. You must identify which of your answers is using a correlated subquery and what about the subquery makes it a correlated subquery. **For a subquery to be correlated a requirement is that it must be the case that the subquery has to be reevaluated for each row of the virtual table defined by the FROM clause of the outer query. A necessary condition for this property is that the correlation variable defined in the outer query is used by including an attribute name.** An example would be `T.course_id` if the correlation variable is named T and the University database schema of our textbook was being used. See Subsection 3.8.3 of our textbook for some examples of what correlated subqueries look like.
- On the course website just below the `assg3_tester.sql` link is a link to a tar file containing an example database instance that you can use for testing. The database instance is in the format that the `psql \copy` command expects for doing bulk loading into a table in a relational database. Each line represents one record (tuple) with the fields in the line representing the attributes in the tuple and are separated by a tab character. There are no blank lines in each data file since a blank line causes a syntax error by the `\copy` command.
This example database instance is the one that the `assg3_tester.sql` refers to when saying that the result of a query should be a particular set of values. To extract the tar file the command is

```
tar xvf sail_db_data.tar
```

in the directory in which you want the subdirectory containing the data files to appear.

Administrative

Your score on this assignment is based on well your batch files run on agora using postgresql. I will run them as my username and database.

You can work in teams of one or two students. If you are in a team of two both of your names must be listed in the readme file. You can talk with people outside your team about the general concepts involved but they cannot help you and you cannot help them with how to answer the questions. Use `handin` on agora using the command

```
handin.453.1 3 assg3.tar
```

to turn in your assignment. If you are in the directory containing the `assg3` directory as a subdirectory, then the command

```
tar cvf assg3.tar assg3
```

will create the `assg3.tar` file in the current directory. The tar file must contain

- the `psql` batch file for dropping all your tables
- the `psql` batch file create all your tables

- the psql batch file for populating all your tables
- the psql batch file for the select statements for your queries
- your readme file

The assignment is due at 5 pm on Wednesday, the 29th of March. No late assignments will be accepted.