Lab 3 - SQL

Big Data Spring 2016

February 8, 2016

Review: Basic SQL Queries

- A basic SQL query has the form

```
SELECT      [DISTINCT] target-list
FROM        relation-list
WHERE       qualification
```

- target-list: a list of attributes of relations in relation-list
- relation-list: a list of relation names (possibly with correlation name)
- qualification: comparisons using defined operators (e.g., >, <, =), which can be combined using AND, OR, and NOT
- DISTINCT: an optional keyword indicating that answer should not contain duplicates

Our Example Today

- 3 tables: sailors, boats, reserves
- Step 1: Create the tables and populate them.
  - We have given you a script to do this, sailors-mysql.sql
  - Type “source path_to_file/sailors-mysql.sql”, where path_to_file is the path to where you’ve saved this file on your laptop

```
create table sailors(
        sid int PRIMARY KEY,
        sname varchar(30),
        rating int,
        age int
);
```

```
create table boats(
        bid int PRIMARY KEY,
        bname char(20),
        color char(10)
);
```

```
create table reserves(
        sid int,
        bid int,
        day date,
        PRIMARY KEY (sid, bid, day)
);
```

Practice Writing SQL Queries

- We have given you a long list of SQL queries to write using this data

- Work on this lab in **groups of 2 or 3** people!

- We will work through a few select queries together
- You can work on the rest for the remainder of the lab (or later at home)

Practice Writing SQL Queries

- 2a) Find the names and ages of all sailors

Practice Writing SQL Queries

- 2a) Find the names and ages of all sailors

```
SELECT sname, age FROM sailors;
```

Practice Writing SQL Queries

- 2a) Find the names and ages of all sailors

```
SELECT sname, age FROM sailors;
```

```
SELECT S.sname, S.age FROM sailors S;
```

Correlation name. Not always necessary, but good practice to use this.

Practice Writing SQL Queries

- 2b) Find all sailors with a rating above 7.

Practice Writing SQL Queries

- 2b) Find all sailors with a rating above 7.

```
SELECT *
FROM sailors
WHERE rating > 7
```

Practice Writing SQL Queries

- 2c) Find the names of sailors who have reserved boat number 103

Practice Writing SQL Queries

- 2c) Find the names of sailors who have reserved boat number 103

```
SELECT sname
FROM sailors S, reserves R
WHERE S.sid = R.sid AND bid = 103
```

Practice Writing SQL Queries

- 2c) Find the names of sailors who have reserved boat number 103

```
SELECT sname
FROM sailors S, reserves R
WHERE S.sid = R.sid AND bid = 103
```

Using a nested query:

```
SELECT sname
FROM sailors S
WHERE S.sid in (SELECT R.sid
                FROM reserves R
                WHERE R.bid = 103)
```

Practice Writing SQL Queries

- 2d) Find the sids of sailors who have reserved a red boat.

Practice Writing SQL Queries

- 2d) Find the sids of sailors who have reserved a red boat.

```
SELECT sid
FROM  reserves R, boats B
WHERE R.bid = B.bid AND color = 'red'
```

Practice Writing SQL Queries

- 2d) Find the sids of sailors who have reserved a red boat.

```
SELECT sid
FROM  reserves R, boats B
WHERE R.bid = B.bid AND color = 'red'
```

This contains duplicates. To remove duplicates, use DISTINCT keyword:

```
SELECT DISTINCT sid
FROM  reserves R, boats B
WHERE R.bid = B.bid AND color = 'red'
```

Practice Writing SQL Queries

- 2h) Find the names of sailors who have reserved both a red and a green boat.

Practice Writing SQL Queries

- 2h) Find the names of sailors who have reserved both a red and a green boat.

Here is one incorrect query:

```
SELECT sname
FROM sailors S, reserves R, boats B
WHERE S.sid = R.sid AND R.bid = B.bid AND (color =
'red' AND color = 'green')
```

What happens?

Practice Writing SQL Queries

- 2h) Find the names of sailors who have reserved both a red and a green boat.

Here is one incorrect query:

```
SELECT sname
FROM sailors S, reserves R, boats B
WHERE S.sid = R.sid AND R.bid = B.bid AND (color =
'red' AND color = 'green')
```

What happens?

Another mistake: a sailor named Horatio has reserved a red boat, and a *different* sailor named Horatio has reserved a green boat – make sure to write your query such that Horatio is not returned as a sailor that has reserved both a red and green boat!

Practice Writing SQL Queries

- 2h) Find the names of sailors who have reserved both a red and a green boat.

Here is one example of a correct query:

```
SELECT DISTINCT S.sname
FROM sailors S, boats B, reserves R
WHERE S.sid = R.sid AND R.bid = B.bid AND
B.color='red' AND S.sid IN
      (SELECT S2.sid
      FROM sailors S2, boats B2, reserves R2
      WHERE S2.sid=R2.sid AND R2.bid=B2.bid AND
      B2.color='green')
```

Practice Writing SQL Queries

- 2j) Find the names of sailors who have not reserved boat number 103.

Practice Writing SQL Queries

- 2j) Find the names of sailors who have not reserved boat number 103.

```
SELECT sname
FROM sailors S
WHERE S.sid NOT IN
      (SELECT sid
       FROM Reserves
      WHERE bid = 103)
```

Practice Writing SQL Queries

- 2k) Find the names of sailors whose rating is better than some sailor called Horatio.

Practice Writing SQL Queries

- 2k) Find the names of sailors whose rating is better than some sailor called Horatio.

```
SELECT sname
FROM sailors S1
WHERE S1.rating > ANY (SELECT rating
                       FROM sailors S2
                       WHERE sname='horatio')
```

Practice Writing SQL Queries

- 2n) Find the average age of sailors with a rating of 10.

Practice Writing SQL Queries

- 2n) Find the average age of sailors with a rating of 10.

```
SELECT AVG(age)
FROM Sailors
WHERE rating = 10
```

Practice Writing SQL Queries

- 2o) Find the name and age of the oldest sailor.

Practice Writing SQL Queries

- 2o) Find the name and age of the oldest sailor.

  What happens if we write the query:

```
SELECT sname,MAX(age)
FROM sailors
```

Practice Writing SQL Queries

- 2o) Find the name and age of the oldest sailor.

  What happens if we write the query:

```
SELECT sname,MAX(age)
FROM sailors
```

  The answer is incorrect! sname is neither in an aggregate nor in GROUP BY

Practice Writing SQL Queries

- 2o) Find the name and age of the oldest sailor.

A correct query:

```
SELECT S.sname, S.age
FROM sailors S
WHERE S.age = (SELECT MAX(S2.age) FROM sailors S2);
```

Does this match the age that was input by the script?

Can you see why?