Announcements

- Midterm exam in CMA 2.306 on 10/29.
- Review session next class.
1) How many records are produced by Q1 when run on the ACL tables shown?

Q1: SELECT * FROM ACL_Artist_2017 UNION
    SELECT * FROM ACL_Artist_2018;

A. 6
B. 10
C. 11
D. 12
2) How many records are produced by Q2 when run on the ACL tables shown?

Q2: SELECT * FROM ACL_Artist_2017 UNION SELECT id FROM ACL_Artist_2018;

<table>
<thead>
<tr>
<th>ACL_Artist_2017</th>
<th>ACL_Artist_2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>id</td>
<td>artist_name</td>
</tr>
<tr>
<td>jz</td>
<td>Jay Z</td>
</tr>
<tr>
<td>sp</td>
<td>Spoon</td>
</tr>
<tr>
<td>hcp</td>
<td>Red Hot Chili Peppers</td>
</tr>
<tr>
<td>atw</td>
<td>Asleep At The Wheel</td>
</tr>
<tr>
<td>ra</td>
<td>Ryan Adams</td>
</tr>
<tr>
<td>sor</td>
<td>School of Rock</td>
</tr>
</tbody>
</table>

A. 0  
B. 6  
C. 10  
D. 12
3) The queries Q1 and Q2 are functionally equivalent.

Q1: SELECT id, artist_name FROM ACL_Artist_2017 UNION
    SELECT id, artist_name FROM ACL_Artist_2018;

Q2: SELECT a.id, a.artist_name
    FROM ACL_Artist_2017 a
    LEFT JOIN ACL_Artist_2018 b ON a.id = b.id
    AND a.artist_name = b.artist_name;

A. True       B. False
4) How many records are produced by Q4 when run on the ACL tables shown?

Q4: SELECT * FROM ACL_Artist_2017 INTERSECT SELECT * FROM ACL_Artist_2018;

A. 2  
B. 4  
C. 6  
D. 10
5) The queries $Q_1$ and $Q_2$ are functionally equivalent.

$Q_1$: SELECT id, artist_name FROM ACL_Artist_2017 INTERSECT
    SELECT id, artist_name FROM ACL_Artist_2018;

$Q_2$: SELECT a.id, a.artist_name
    FROM ACL_Artist_2017 a
    JOIN ACL_Artist_2018 b ON a.id = b.id
    AND a.artist_name = b.artist_name;

A. True    B. False
What’s wrong with this design?

- Insert Anomaly
- Update Anomaly
- Delete Anomaly
Design Principles

- one entity type per table
- each table has a primary key
- referential integrity
- data types represent domain of values
Normal Forms

1NF: A database schema is in 1NF iff all attributes have scalar values.

2NF: 1NF + all non-key attributes must be functionally determined by the entire primary key.

3NF: 2NF + all non-key attributes must be functionally determined by only the primary key.

Functional Dependencies:
If two records agree on the attributes $A_1, A_2, \ldots, A_n$ then they must also agree on the attributes $B_1, B_2, \ldots, B_n$

Formally:
$A_1, A_2, \ldots, A_n \rightarrow B_1, B_2, \ldots, B_n$
Normal Form Violations

Teacher
- PK tid fname lname dept salary

Teaches
- PK, FK tid cno cname

Student
- PK sid fname lname dob email

Takes
- PK, FK sid cno grade GPA

Class
- PK cno cname credits semester
BigQuery Demo
Practice Problem

Formulate a SQL query that finds any orphan sid records in the table Takes.

Student(sid, fname, Iname, dob)
Class(cno, cname, credits)
Teacher(tid, fname, Iname, dept)
Takes(sid, cno, grade)
Teaches(tid, cno)
iClicker Question

Formulate a SQL query that finds any orphan sid records in the table Takes.

Does the query require an outer join?

A. Yes
B. No