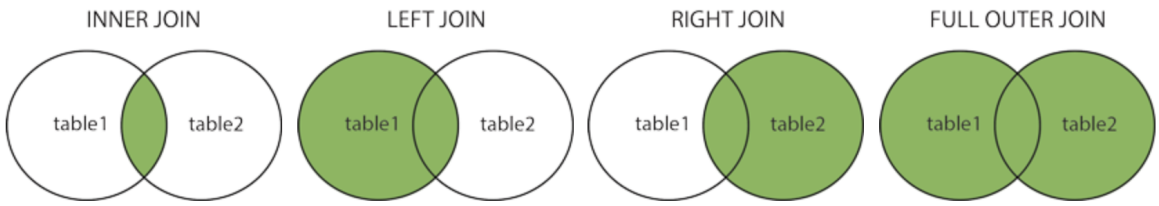
DS 100: Principles and Techniques of Data Science **Date: November 7, 2018**

Discussion #10

Name:

SQL Joins



Note: You do not always have to use the JOIN keyword to join sql tables. The following are equivalent:

```
SELECT column1, column2
FROM table1, table2
WHERE table1.id = table2.id;

SELECT column1, column2
FROM table1 JOIN table2
ON table1.id = table2.id;
```

1. Describe which records are returned from each type of join.

SQL

2. Circle TRUE or FALSE.
   (a) True False SQL is a declarative language that specifies what to produce but not how to compute it.
   (b) True False The primary key of a relation is the column or set of columns that determine the values of the remaining column.
   (c) True False The schema of a table consists of the data stored in the table.
   (d) True False The WHERE and HAVING clause can be used interchangeably as they perform the same operation.

Writing SQL Queries

Consider the following schema:

```
Clowns(cid int, cname text, booth text)
Balloons(bid int, bshape text, bcolor text)
Catalog(cid int, bid int, cost float)
```

Note: The Catalog table contains prices for Balloons sold by different Clowns standing at certain booths in a fair.

3. How may we query for the top 3 most expensive shapes sold by Whompers LeFou, ignoring the possibility that Whompers could be selling the same shape in different colors?

4. How may we query for the top 3 most expensive shapes sold by Whompers LeFou, taking into consideration the possibility that Whompers could be selling the same shape in different colors by using the highest-priced color of each shape?

5. What is the average cost of a red balloon at booths that offer more than 3 red shapes per clown? Note that each clown at the booth does not necessarily have to be selling more than 3 shapes.

6. Consider the following real estate schema:

```
Homes(home_id int, city text, bedrooms int, bathrooms int,
area int)
Transactions(home_id int, buyer_id int, seller_id int,
transaction_date date, sale_price int)
Buyers(buyer_id int, name text)
Sellers(seller_id int, name text)
```

For the query language questions below, fill in the blanks in the answer to complete the query. For each SQL query and nested subquery, please start a new line when you reach a SQL keyword (SELECT, WHERE, AND, etc.). However, do not start a new line for aggregate functions (COUNT, SUM, etc.), and comparisons (LIKE, AS, IN, NOT IN, EXISTS, NOT EXISTS, ANY, or ALL.)

(a) Fill in the blanks in the SQL query to find the duplicate-free set of id's of all homes in Berkeley with at least 6 bedrooms and at least 2 bathrooms that were bought by "Bobby Tables."

```
SELECT _____
FROM _____
WHERE _____
_____
_____
_____
_____
_____
```

(b) Fill in the blanks in the SQL query to find the id and selling price for each home in Berkeley. If the home has not ben sold yet, **the price should be NULL**.

```
SELECT _____
FROM _____
_____ JOIN _____
ON _____
WHERE _____;
```