Null Values: Solutions

1. Suppose we have a table called Runnymede with the following content:

name		age	1	grade
diane				8
will				8
cate				1
tom				
micah				1
jamieson				2
(6 rows)				

What is the output of each query below?

- (a) select min(grade), max(grade), sum(grade), avg(grade), count(grade), count(*)
 from Runnymede;
- (b) select min(age), max(age), sum(age), avg(age), count(age), count(*)
 from Runnymede;

Solution:

```
csc343h-dianeh=> select min(grade), max(grade), sum(grade), avg(grade),
csc343h-dianeh-> count(grade), count(*)
csc343h-dianeh-> from Runnymede;
min | max | sum |
                                | count | count
----+----+-----+-----
       8 | 20 | 4.0000000000000000000 |
                                    5 l
  1 |
(1 row)
csc343h-dianeh=> -- In the age column, *every* value is null!
csc343h-dianeh=> select min(age), max(age), sum(age), avg(age),
csc343h-dianeh-> count(age), count(*)
csc343h-dianeh-> from Runnymede;
min | max | sum | avg | count | count
        1
(1 row)
```

Aggregation ignores nulls. Notice that count(age) can report a value even though there is nothing but nulls in the age column.

2. We have tables R and T. Their contents are shown below.

```
R T
a | b | b | c
---+--
1 | 2 2 | 5
```

What is the result of this query:

```
select * from R natural join T;
```

Solution:

Why are tuples (5, null) and (null, 18) excluded from the result, yet tuples (null, 6) and (6, 88) are included? (5, null) and (null, 18) are excluded because when the natural join compares the 2 null values, they are not considered equal. It makes sense if you think of natural join as involving a Cartesian product and a WHERE, and if you remember that the truth value of a comparison involving nulls is always unknown and that WHERE is picky. In contrast, (null, 6) and (6, 88) are included because we are comparing on attribute b, which is 6 for both of these tuples, not null.