1. Question ID 557: The table Arc(x,y) currently has the following tuples (note there are duplicates): (1,2), (1,2), (2,3), (3,4), (3,4), (4,1), (4,1), (4,1), (4,2). Compute the result of the query:

```sql
SELECT a1.x, a2.y, COUNT(*)
FROM Arc a1, Arc a2
WHERE a1.y = a2.x
GROUP BY a1.x, a2.y;
```

Question Explanation: The query is asking for the number of paths of length 2 between each pair of nodes a1.x and a2.y. To compute these numbers, we can theta-join Arc with itself, with the condition that the second component of the first copy equal the first component of the second copy. Here is a list of all the tuples and their multiplicities that we get:

- (1,2,2,3) [twice]
- (2,3,3,4) [twice]
- (3,4,4,1) [6 times]
- (3,4,4,2) [twice]
- (4,1,1,2) [6 times]
- (4,2,2,3) [once]

Since no two of these tuples have the same first and 4th components, we do not have to combine any of these six groups to get the grouping according to the SQL GROUP-BY. Thus, the tuples of the result are formed from the first and fourth components and the multiplicity of the tuple.

2. Question ID 577: The table

Scores(Team, Day, Opponent, Runs)

Gives the scores in the Japanese Baseball League for two consecutive days. The data in this table is as follows:

Team	Day	Opponent	Runs
Dragons	Sunday	Swallows	4
Tigers	Sunday	Bay Stars	9
Carp	Sunday	Giants	2
Swallows	Sunday	Dragons	7
Bay Stars	Sunday	Tigers	2
Giants	Sunday	Carp	4

Dragons	Monday	Carp	6
Tigers	Monday	Bay Stars	5
Carp	Monday	Dragons	3
Swallows	Monday	Giants	0
Bay Stars	Monday	Tigers	7
Giants	Monday	Swallows	5

Determine the result of the query

```
SELECT Team, SUM(Runs)
FROM Scores
GROUP BY Team
```

Question Explanation: Each team played two games, so the team appears in the result next to the sum of the numbers of runs they scored in those two games. The result is:

Team	SUM(Runs)
Dragons	10
Tigers	14
Carp	5
Swallows	7
Bay Stars	9
Giants	9

3. Question ID 580: Suppose we have a relation with schema
   ```
   R(A, B, C, D, E)
   ```

If we issue a query of the form

```
SELECT ???
FROM R
WHERE ...
GROUP BY C, D
```

What terms can appear in the SELECT list (represented by ??? in the above query)?

Question Explanation: When there is a GROUP-BY clause, any aggregated term can appear in the SELECT list. However, an attribute that is not in the GROUP-BY list cannot appear, unaggregated, in the

SELECT list. Thus, C or D may appear unaggregated, and all five attributes can appear in an aggregation. However, A, B, or E cannot appear by themselves.

4. Question ID 583: Suppose we have a relation with schema

```
R(A, B, C, D, E)
```

If we issue a query of the form

```
SELECT ...
FROM R
WHERE ...
GROUP BY B, E
HAVING ???
```

What terms can appear in the HAVING condition (represented by ??? in the above query)?.

Question Explanation: Any aggregated term can appear in the HAVING condition. However, an attribute that is not in the GROUP-BY list cannot appear, unaggregated, in the HAVING condition. Thus, B or E may appear unaggregated, and all five attributes can appear in an aggregation. However, A, C, or D cannot appear by themselves.

5. Question ID 593: Here are three relations, R(A,B), S(C,D), and T(E,F). Their current values are:

R		S		T	
A	**B**	**C**	**D**	**E**	**F**
0	1	0	1	0	1
1	0	1	0	1	0
1	1	1	1	1	1

Compute the result of the query:

```
SELECT A, F, SUM(C), SUM(D)
FROM R, S, T
WHERE B = C AND D = E
GROUP BY A, F
HAVING COUNT(*) > 1
```

Question Explanation: Row (1,1) of S joins with two rows --- (0,1) and (1,1) --- of R and with two rows --- (1,0) and (1,1) --- of T, yielding four rows in the result. Row (0,1) of S joins with only (1,0) of R, but two rows of T: (1,0) and (1,1). Likewise, (1,0) if S joins with two rows of R --- (0,1) and (1,1) --- but only one row of T: (0,1). There are thus eight rows in the result of the select-from-where part. We show these eight rows below, grouped by their A and F values.

A	B	C	D	E	F
0	1	1	1	1	0

0	1	1	1	1	1
0	1	1	0	0	1

1	1	1	1	1	0
1	0	0	1	1	0

1	1	1	1	1	1
1	0	0	1	1	1
1	1	1	0	0	1

Finally, we must apply the HAVING clause. That clause says the group must have more than one row. Therefore, the first group, corresponding to A=F=0, is eliminated. However, the remaining three groups survive. The three resulting rows are:

A	F	SUM(C)	SUM(D)
0	1	2	1
1	0	1	2
1	1	2	2