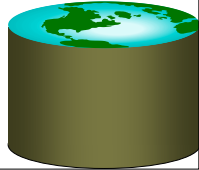


SQL: The Query Language Part 2

Lecture 10
R & G - Chapter 5

The important thing is not to stop questioning.

Albert Einstein



Example Instances

Reserves

sid	bid	day
22	101	10/10/96
95	103	11/12/96

Sailors

sid	sname	rating	age
22	Dustin	7	45.0
31	Lubber	8	55.5
95	Bob	3	63.5

Boats

bid	bname	color
101	Interlake	blue
102	Interlake	red
103	Clipper	green
104	Marine	red

Queries With GROUP BY

- To generate values for a column based on groups of rows, use **aggregate functions** in **SELECT** statements with the **GROUP BY** clause

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

The **target-list** contains (i) list of column names & (ii) terms with aggregate operations (e.g., **MIN (S.age)**).

- column name list (i) can contain only attributes from the **grouping-list**.

Group By Examples

For each rating, find the average age of the sailors

```
SELECT S.rating, AVG (S.age)
FROM Sailors S
GROUP BY S.rating
```

For each rating find the age of the youngest sailor with age ≥ 18

```
SELECT S.rating, MIN (S.age)
FROM Sailors S
WHERE S.age >= 18
GROUP BY S.rating
```

Conceptual Evaluation

- The cross-product of **relation-list** is computed, tuples that fail **qualification** are discarded, 'unnecessary' fields are deleted, and the remaining tuples are partitioned into groups by the value of attributes in **grouping-list**.
- One answer tuple is generated per qualifying group.

```
SELECT S.rating, MIN (S.age)
FROM Sailors S
WHERE S.age >= 18
GROUP BY S.rating
```

Answer Table

rating	age
1	33.0
7	35.0
8	55.0
10	35.0

3. Perform Aggregation

1. Form cross product

2. Delete unneeded columns, rows; form groups



Find the number of reservations for each red boat.

```
SELECT B.bid, COUNT(*) AS numres
FROM Boats B, Reserves R
WHERE R.bid=B.bid
      AND B.color='red'
GROUP BY B.bid
```

- Grouping over a join of two relations.



```
SELECT B.bid, COUNT (*) AS scout
FROM Boats B, Reserves R
WHERE R.bid=B.bid AND B.color='red'
GROUP BY B.bid
```

b.bid	b.color	r.bid
101	blue	101
102	red	101
103	green	101
104	red	101
101	blue	102
102	red	102
103	green	102
104	red	102

b.bid	b.color	r.bid
102	red	102

b.bid	scout
102	1

answer



Queries With GROUP BY and HAVING

```
SELECT [DISTINCT] target-list
FROM relation-list
WHERE qualification
GROUP BY grouping-list
HAVING group-qualification
```

- Use the HAVING clause with the GROUP BY clause to restrict which group-rows are returned in the result set



Conceptual Evaluation

- Form groups as before.
- The *group-qualification* is then applied to eliminate some groups.
 - Expressions in *group-qualification* must have a *single value per group!*
 - That is, attributes in *group-qualification* must be arguments of an aggregate op or must also appear in the *grouping-list*. (SQL does not exploit primary key semantics here!)
- One answer tuple is generated per qualifying group.



Find the age of the youngest sailor with age >= 18, for each rating with at least 2 such sailors

```
SELECT S.rating, MIN (S.age)
FROM Sailors S
WHERE S.age >= 18
GROUP BY S.rating
HAVING COUNT (*) > 1
```

sid	sname	rating	age
22	Dustin	7	45.0
31	lubber	8	55.5
71	zorba	10	16.0
64	horatio	7	35.0
29	brutus	1	33.0
58	rusty	10	35.0

rating	age
1	33.0
7	45.0
7	35.0
8	55.5
10	35.0

rating	m-age	count
1	33.0	1
7	35.0	2
8	55.0	1
10	35.0	1

rating	age
7	35.0

Answer relation



Find sailors who've reserved all boats.

```
SELECT S.sname
FROM Sailors S
WHERE NOT EXISTS (SELECT B.bid
                  FROM Boats B
                  WHERE NOT EXISTS (SELECT R.bid
                                    FROM Reserves R
                                    WHERE R.bid=B.bid
                                       AND R.sid=S.sid))
```



Find sailors who've reserved all boats.

- Can you do this using Group By and Having?

```
SELECT S.name
FROM Sailors S, reserves R
WHERE S.sid = R.sid
GROUP BY S.name, S.sid
HAVING
COUNT(DISTINCT R.bid) =
  (Select COUNT (*) FROM Boats)
```

Note: must have both sid and name in the GROUP BY clause. Why?



```
SELECT S.name, S.sid
FROM Sailors S, reserves R
WHERE S.sid = r.sid
GROUP BY S.name, S.sid
HAVING
COUNT(DISTINCT R.bid) =
  Select COUNT (*) FROM Boats
```

s.name	s.sid	r.sid	r.bid
Dustin	22	22	101
Lubber	95	22	101
Bob	95	22	101
Dustin	22	95	102
Lubber	95	95	102
Bob	95	95	102

bid	bname	color
101	Interlake	blue
102	Interlake	red
103	Clipper	green
104	Marine	red

Count (*) from boats = 4

s.name	s.sid	bcount
Dustin	22	1
Bob	95	1

Apply having clause to groups

s.name	s.sid



INSERT

```
INSERT [INTO] table_name [(column_list)]
VALUES (value_list)
```

```
INSERT [INTO] table_name [(column_list)]
<select statement>
```

```
INSERT INTO Boats VALUES ( 105, 'Clipper', 'purple')
INSERT INTO Boats (bid, color) VALUES (99, 'yellow')
```

You can also do a "bulk insert" of values from one table into another:

```
INSERT INTO TEMP(bid)
SELECT r.bid FROM Reserves R WHERE r.sid = 22;
(must be type compatible)
```



DELETE & UPDATE

```
DELETE [FROM] table_name
[WHERE qualification]
```

```
DELETE FROM Boats WHERE color = 'red'
```

```
DELETE FROM Boats b
WHERE b.bid =
  (SELECT r.bid FROM Reserves R WHERE r.sid = 22)
```

Can also modify tuples using UPDATE statement.

```
UPDATE Boats
SET Color = "green"
WHERE bid = 103;
```



Null Values

- Field values in a tuple are sometimes **unknown** (e.g., a rating has not been assigned) or **inapplicable** (e.g., no spouse's name).
 - SQL provides a special value *null* for such situations.
- The presence of **null** complicates many issues. E.g.:
 - Special operators needed to check if value is/is not *null*.
 - Is *rating* > 8 true or false when *rating* is equal to *null*? What about AND, OR and NOT connectives?
 - We need a **3-valued logic** (true, false and *unknown*).
 - Meaning of constructs must be defined carefully. (e.g., WHERE clause eliminates rows that don't evaluate to true.)
 - New operators (in particular, *outer joins*) possible/needed.



Joins

```
SELECT (column_list)
FROM table_name
[INNER | {LEFT | RIGHT | FULL} OUTER] JOIN table_name
ON qualification_list
WHERE ...
```

Explicit join semantics needed unless it is an INNER join (INNER is default)



Inner Join

Only the rows that match the search conditions are returned.

```
SELECT s.sid, s.name, r.bid
FROM Sailors s INNER JOIN Reserves r
ON s.sid = r.sid
```

Returns only those sailors who have reserved boats
SQL-92 also allows:

```
SELECT s.sid, s.name, r.bid
FROM Sailors s NATURAL JOIN Reserves r
```

"NATURAL" means equi-join for each pair of attributes with the same name



```
SELECT s.sid, s.name, r.bid
FROM Sailors s INNER JOIN Reserves r
ON s.sid = r.sid
```

sid	sname	rating	age	sid	bid	day
22	Dustin	7	45.0	22	101	10/10/96
31	Lubber	8	55.5	95	103	11/12/96
95	Bob	3	63.5			

s.sid	s.name	r.bid
22	Dustin	101
95	Bob	103



Left Outer Join

Left Outer Join returns all matched rows, plus all unmatched rows from the table on the left of the join clause

(use nulls in fields of non-matching tuples)

```
SELECT s.sid, s.name, r.bid
FROM Sailors s LEFT OUTER JOIN Reserves r
ON s.sid = r.sid
```

Returns all sailors & information on whether they have reserved boats



```
SELECT s.sid, s.name, r.bid
FROM Sailors s LEFT OUTER JOIN Reserves r
ON s.sid = r.sid
```

sid	sname	rating	age	sid	bid	day
22	Dustin	7	45.0	22	101	10/10/96
31	Lubber	8	55.5	95	103	11/12/96
95	Bob	3	63.5			

s.sid	s.name	r.bid
22	Dustin	101
95	Bob	103
31	Lubber	



Right Outer Join

Right Outer Join returns all matched rows, plus all unmatched rows from the table on the right of the join clause

```
SELECT r.sid, b.bid, b.name
FROM Reserves r RIGHT OUTER JOIN Boats b
ON r.bid = b.bid
```

Returns all boats & information on which ones are reserved.



```
SELECT r.sid, b.bid, b.name
FROM Reserves r RIGHT OUTER JOIN Boats b
ON r.bid = b.bid
```

sid	bid	day	bid	bname	color
22	101	10/10/96	101	Interlake	blue
95	103	11/12/96	102	Interlake	red
			103	Clipper	green
			104	Marine	red

r.sid	b.bid	b.name
	22	101 Interlake
		102 Interlake
	95	103 Clipper
		104 Marine



Full Outer Join

Full Outer Join returns all (matched or unmatched) rows from the tables on both sides of the join clause

```
SELECT r.sid, b.bid, b.name
FROM Reserves r FULL OUTER JOIN Boats b
ON r.bid = b.bid
```

Returns all boats & all information on reservations



```
SELECT r.sid, b.bid, b.name
FROM Reserves r FULL OUTER JOIN Boats b
ON r.bid = b.bid
```

sid	bid	day	bid	bname	color
22	101	10/10/96	101	Interlake	blue
95	103	11/12/96	102	Interlake	red
			103	Clipper	green
			104	Marine	red

r.sid	b.bid	b.name
22	101	Interlake
	102	Interlake
95	103	Clipper
	104	Marine

Note: in this case it is the same as the ROJ because bid is a foreign key in reserves, so all reservations must have a corresponding tuple in boats.



Views

```
CREATE VIEW view_name
AS select_statement
```

Makes development simpler
Often used for security
Not instantiated - makes updates tricky

```
CREATE VIEW Reds
AS SELECT B.bid, COUNT (*) AS scount
FROM Boats B, Reserves R
WHERE R.bid=B.bid AND B.color='red'
GROUP BY B.bid
```



```
CREATE VIEW Reds
AS SELECT B.bid, COUNT (*) AS scount
FROM Boats B, Reserves R
WHERE R.bid=B.bid AND B.color='red'
GROUP BY B.bid
```

b.bid	scount	Reds
102	1	