MySQL RDBMS

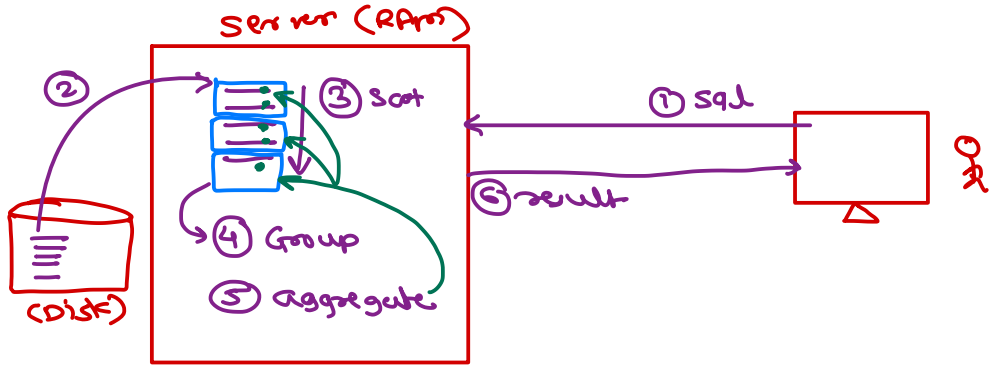
Trainer: Mr. Nilesh Ghule

GROUP BY clause

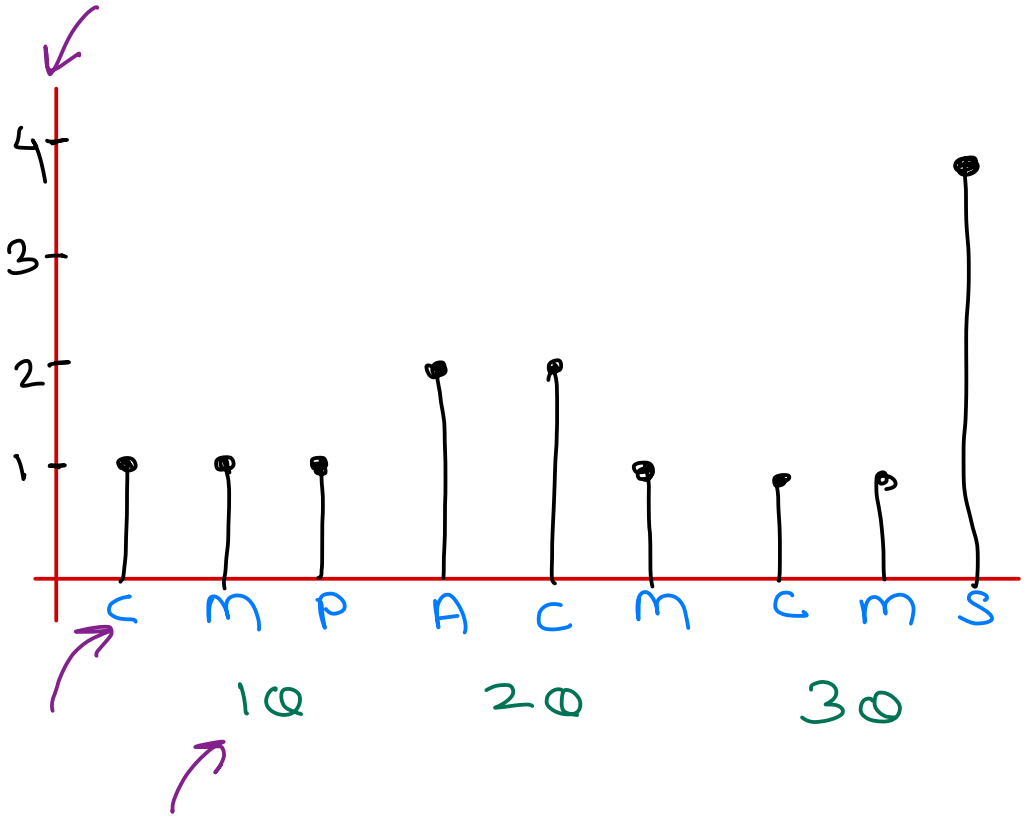
- GROUP BY is used for analysis of data i.e. generating reports & charts.
- When GROUP BY single column, generated output can be used to plot 2-D chart. When GROUP BY two column, generated output can be used to plot 3-D chart and so on.
- GROUP BY queries are also called as Multi-dimensional / Spatial queries.
- Syntactical Characteristics:
  - If a column is used for GROUP BY, then it may or may not be used in SELECT clause.
  - If a column is in SELECT, it must be in GROUP BY.

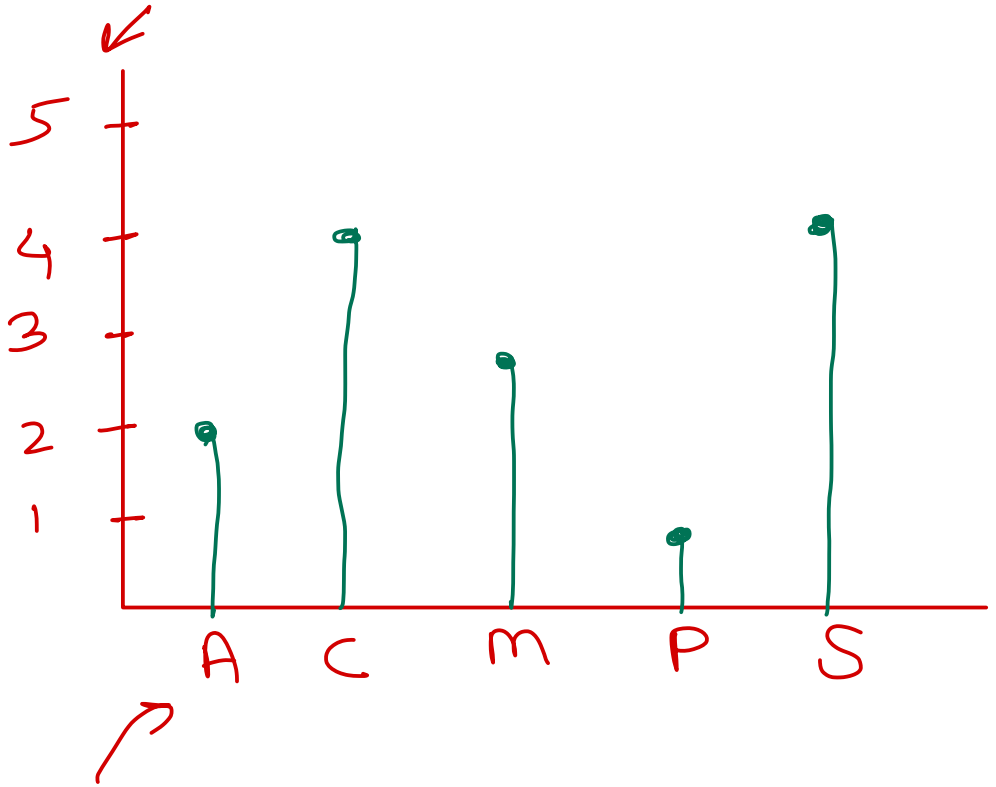
```
select job, count(empno) from emp group by job
having job in (M, A);
```

- When GROUP BY query is fired on database server, it does following:
  - Load data from server disk into server RAM.
  - Sort data on group by columns.
  - Group similar records by group columns.
  - Perform given aggregate ops on each column.
  - Send result to client.




```
select job, count(empno) from emp where job in (M, A)
group by job;
```


4
3
2
1
C M P A C M C M S
100
200
300


5
4
3
2
1
A C M P S

HAVING clause

- HAVING clause cannot be used without GROUP BY clause.
- HAVING clause is used to specify condition on aggregate values/fns.
- Examples:
  - SELECT deptno, SUM(sal) FROM EMP GROUP BY deptno HAVING SUM(sal) > 9000;
- Syntactical Characteristics:
  - WHERE clause executed for each record; while HAVING is executed for each group.
  - HAVING clause can be used to specify condition on group fn or grouped columns.
  - However using HAVING to specify condition of group col reduce the performance. Use WHERE clause for the same.
- Examples:
  - SELECT deptno, SUM(sal) FROM EMP GROUP BY deptno HAVING deptno = 20;
  - SELECT deptno, SUM(sal) FROM EMP WHERE deptno = 20 GROUP BY deptno; ← efficient.
- We may use GROUP BY with WHERE, ORDER BY & LIMIT.

Thank you!

Nilesh Ghule <nilesh@sunbeaminfo.com>