SQL: Structured Query Language
Grouping Queries

SQL SELECT statement has two more clauses to support grouping operations: GROUP BY and HAVING clauses.

GROUP BY Clause

The syntax of a GROUP BY clause is

\[ \text{GROUP BY } \text{AttributeName}, ..., \text{AttributeName} \]

The GROUP BY clause is added to the SELECT statement after the WHERE clause (or, if there is no WHERE clause, after the FROM clause.

GROUP BY clause causes the DBMS to separate the cartesian product of all tables referenced in the FROM clause into groups of tuples.

Each group of tuples must agree on all values of attributes listed in the GROUP BY clause.

Grouping operation is used to allow for computation and reporting of aggregate operations over groups in the SELECT statement.

Consider, for example the relational table

Student(FirstName, LastName, GPA, Class, Grade, School)

The following query

\[
\text{SELECT School, COUNT(*)}
\text{FROM Student}
\text{GROUP BY School;}
\]

will output the number of students enrolled in each school.

The following rules need to be observed concerning the content of the SELECT clause. If a GROUP BY clause appears in the SELECT statement, then the SELECT can contain only the following:
Attributes that are listed in the \texttt{GROUP BY} clause.

- Aggregate operations on attributes not listed in the \texttt{GROUP BY} clause.
- \texttt{COUNT(\ast)}

**HAVING Clause**

\texttt{GROUP BY} clause transforms the cartesian product from a relation of tuples into a relation of groups of tuples.

\texttt{HAVING} clause to the groups is what \texttt{WHERE} clause to the individual tuples. It provides a condition, and filters out the groups that fail it.

The syntax of \texttt{HAVING} clause is:

\texttt{HAVING <Condition>}

Here, the condition is a boolean combination of conditions that involve:

- Attributes from the \texttt{GROUP BY} list
- Aggregate expressions over attributes \texttt{not} from the \texttt{GROUP BY} list.

For example

\begin{verbatim}
SELECT School, COUNT(\ast) 
FROM Student 
GROUP BY School 
HAVING AVG(GPA) > 3.0;
\end{verbatim}

returns the number of students enrolled in each school, which has average student GPA over 3.0.