Introduction

In relational databases, the `GROUP BY` and `HAVING` clauses are used to group rows that have the same values in specified columns and to filter grouped records. These clauses are crucial in SQL for performing aggregate calculations and making data analysis efficient and organized.

Objectives

- To understand the purpose and syntax of the `GROUP BY` clause.
- To learn how to apply aggregate functions (`COUNT`, `SUM`, `AVG`, `MIN`, `MAX`) within `GROUP BY`.
- To explore the use of the `HAVING` clause for filtering grouped data.

Theory

1. **GROUP BY Clause**
   The `GROUP BY` clause is used to arrange identical data into groups. It is often used with aggregate functions to perform calculations on each group.

   Syntax:

```sql
SELECT column1, aggregate_function(column2)
FROM table_name
GROUP BY column1;
```

2. **HAVING Clause**

   The `HAVING` clause is used to filter records after they have been grouped. It is similar to the `WHERE` clause, but `HAVING` works on grouped data.

   Syntax:

```sql
SELECT department, COUNT(employee_id) AS employee_count
FROM employees
GROUP BY department;
```

Practical Examples

1. **Basic GROUP BY Example**

   Find the total sales for each product:

```sql
SELECT column1, aggregate_function(column2)
FROM table_name
GROUP BY column1
HAVING condition;
```

2. **GROUP BY with HAVING**

   Find products with total sales greater than 5000:

```sql
SELECT product_name, SUM(sales_amount) AS total_sales
FROM sales
GROUP BY product_name
HAVING SUM(sales_amount) > 5000;
```

3. **Multiple Columns in GROUP BY**

   Find the total sales for each product in each region:

```sql
SELECT region, product_name, SUM(sales_amount) AS total_sales
FROM sales
GROUP BY region, product_name;
```

Case Study

Scenario: A company wants to analyze its employee data.
Task: Use the `GROUP BY` and `HAVING` clauses to:

1. Calculate the average salary of employees in each department.
2. Identify departments where the average salary is above 50,000.

```sql
SELECT department, AVG(salary) AS avg_salary
FROM employees
GROUP BY department
HAVING AVG(salary) > 50000;
```

Conclusion

The `GROUP BY` clause is a powerful tool for summarizing and analyzing data. Combined with the `HAVING` clause, it allows users to filter results based on aggregate calculations, enabling more refined data insights. By mastering these clauses, database professionals can handle complex data queries effectively.