

SQL Order of Execution Overview

The syntax order and order of execution of SQL clauses differ. The following steps show the **typical** processing order of an SQL query. Although the typical order is shown below, the SQL Optimizer can adjust the execution order and still return the identical results

1 FROM

Select and Join Tables

The FROM clause is executed first to **retrieve the working data** for the query. If an ON clause is present, the **join is performed**. If you are joining large tables, it's best to limit the rows prior to the join for a more efficient query.

In SAS, a WHERE= data set option can be used on tables in the FROM clause to filter rows prior to the join.

2 WHERE

Filter the Working Data

After the working data is retrieved, the WHERE clause **filters the rows** based on the WHERE expression(s). To include references to column aliases, you must use the **SAS CALCULATED** keyword or **repeat the calculation** in the WHERE clause.

The WHERE clause does not filter aggregated data.

3 GROUP BY

Aggregate the Working Data

The rows remaining in the working data can then be grouped based on **common values**. The GROUP BY clause groups by the specified column(s) and an aggregate function **summarizes the values for the specified group**. It is a best practice to precede aliases with the CALCULATED keyword.

The GROUP BY clause requires a summary function in the SELECT clause to aggregate the data.

4 HAVING

Filter the Aggregated Data

The HAVING clause instructs SQL to filter the data **after the data is summarized** by the GROUP BY clause. In PROC SQL, aliases are accessible in this clause.

5 SELECT

Retrieve the Final Data

After the working data is filtered and data is summarized, the SELECT clause selects the columns to be **returned**, and any remaining computed columns are **calculated**.

6 ORDER BY

Sort the Final Data

After the results have been selected, **sorting** takes place. You can sort by ascending or descending order. Because aliases have been computed, you can reference them in the clause.

If an INTO clause is present, it executes after the ORDER BY clause.

For in-depth information about how your query executes, use the **_TREE** and **_METHOD** PROC SQL options.