

## COURSE GLOSSARY

# Joining Data in SQL

**Alias (table or column alias):** A temporary name given to a table or column in a query using AS (or just a space) to shorten references or rename output columns for readability

**Anti join:** A query pattern (often implemented with WHERE NOT IN or NOT EXISTS) that returns rows from the first table that do not have matching values in the second table

**Column:** A named field in a table that holds values of a specific data type for every record, often used for conditions, selection, and aggregation

**CROSS JOIN:** A join that produces the Cartesian product of two tables, returning every possible combination of rows from the first and second table

**DISTINCT:** A keyword that removes duplicate rows from a query result, returning only unique records

**EXCEPT:** A set operation that returns rows from the first query that do not appear in the second query, effectively subtracting one result set from another

**FROM clause:** The part of an SQL query that lists the tables or subqueries to select data from and defines the left/right context for joins

**Data leader:** Any individual who champions data use, models data-first behaviors, and helps others learn to use data effectively regardless of their formal title

**LOWESS:** A locally weighted scatterplot smoothing technique that fits a smooth curve through points to reveal nonlinear trends, often used to visualize residual patterns in diagnostic plots

**INTERSECT:** A set operation that returns only the rows that appear in both result sets of two queries, with duplicates removed

**Join:** An SQL operation that combines columns from two or more tables by matching rows according to a specified condition so you can query related data together

**Key (primary key):** A column or set of columns that uniquely identifies each row in a table, used to enforce uniqueness and to join tables reliably

**LEFT JOIN:** A join that returns all rows from the left table and the matching rows from the right table, using NULLs for right-table columns when no match exists

**Many-to-many relationship:** A relationship where multiple records in one table relate to multiple records in another, often represented by a join table that links the two

**ON clause:** The part of a JOIN statement that specifies the condition (one or more column comparisons) used to match rows between the joined tables

**One-to-many relationship:** A relationship where a single record in one table is associated with multiple records in another table, such as author-to-books

**One-to-one relationship:** A relationship where each record in one table corresponds to exactly one record in another table, meaning paired uniqueness across the two tables

**RIGHT JOIN:** A join that returns all rows from the right table and the matching rows from the left table, using NULLs for left-table columns when no match exists

**SELF JOIN:** A join where a table is joined with itself, typically using aliases, to compare or relate rows within the same table

**Semi join:** A query pattern (often implemented with WHERE IN or EXISTS) that returns rows from the first table that have matching values in the second table, without adding columns from the second table

**Set operation:** An SQL operation that combines whole result sets (rows) from two queries—such as UNION, INTERSECT, or EXCEPT—rather than merging columns via keys

**Subquery:** A query nested inside another SQL statement (in SELECT, FROM, or WHERE) that provides values or a temporary table used by the outer query

**Table:** A database object that stores data in rows and columns, where each row is a record and each column is a field with a specific data type

**UNION ALL:** A set operation that stacks results from two queries and returns all rows including duplicates, preserving repeated records

**UNION:** A set operation that stacks results from two queries and returns all distinct rows present in either result set, removing duplicates

**USING clause:** A shorthand for JOIN when the join column has the same name in both tables, allowing you to list that column once in parentheses instead of qualifying it with table names

**WHERE clause:** The part of an SQL query that filters rows based on specified conditions, and is commonly used to implement semi joins and anti joins via subqueries