

Lecture Notes – Distinct Queries

Section 1: Lecture Summary

The **DISTINCT** keyword removes duplicate values from query results, ensuring each value appears only once in the output. It is applied to one or more columns in a SELECT statement and is particularly useful when retrieving data that may contain repeated values across multiple rows. The keyword helps identify unique values within a dataset without manually filtering duplicates.

Section 2: Key Concepts and Explanations

Understanding DISTINCT Functionality

The **DISTINCT** keyword works by examining the selected columns and returning only unique combinations of values. When applied to a single column, it eliminates duplicate entries for that column. When applied to multiple columns, it ensures the entire row combination is unique.

Common Use Cases

Identifying unique department assignments: When you need to know which departments have employees, rather than seeing each department repeated for every employee record.

Finding unique salary values: When multiple employees may earn the same salary, DISTINCT shows each salary amount only once, providing a clear picture of the salary structure without redundancy.

Data exploration: DISTINCT helps understand the variety and range of values in a column, making it useful for data quality checks and reporting.

Section 3: Example Code and Use Cases

Example 1: Retrieving Unique Department IDs with Employees

```
SELECT DISTINCT DeptID  
FROM Employees;
```

This query returns each department ID that has at least one employee. Without `DISTINCT`, the result would show the same `DeptID` multiple times for each employee in that department. With `DISTINCT`, each `DeptID` appears only once, showing which departments are represented in the employee roster.

Example 2: Finding All Unique Salary Amounts

```
SELECT DISTINCT Salary  
FROM Employees;
```

This query displays each unique salary amount assigned to employees. If multiple employees earn the same salary, that amount appears only once in the results. This is valuable for understanding the complete salary structure without duplication.

Section 4: Key Takeaways

The **`**DISTINCT**`** keyword is essential for removing redundancy from query results. It applies to the entire selected columns, ensuring unique combinations appear in the output. `DISTINCT` is most effective when you need to identify the variety of values in a dataset rather than individual records. This keyword will be used frequently in more complex queries as your SQL skills advance.