Introduction to Data Analysis using SQL

Pre-requisites

Hope you have gone through the self-learning content for this session on the PRISM portal.

By the End of this Session, You Will:

- Understand the significance of MySQL in data analytics.
- Learn few miscellaneous functions in MySQL that allows us to perform data manipulations – CAST, DATE, CONCAT, REPLACE.
- Change datatype of existing column using CAST function.
- Extract date from datetime expression using the DATE function.
- Combine text columns using the CONCAT function.
- Substitute text within a larger text using the REPLACE function.
- Learn about PRIMARY and FOREIGN keys.
- Understand the importance of Entity Relationship Diagrams.

What Have We Learned So Far?

- Types of DML commands – INSERT, UPDATE, and DELETE.
- INSERT commands helps us in inserting data into a database.
- UPDATE command allows us to update the existing data in a database.
- DELETE command helps us in removing the existing data from a database.
- Fundamentals of SELECT query.
- Use of WHERE clause to filter the subset of rows from a table.
- Use of ORDER BY clause to sort output result set.
- Use of GROUP BY to aggregate numerical data using values from another column.

Pop Quiz

Q. The SELECT query combined with the GROUP BY clause in SQL is primarily used for _____.

a. Filtering rows based on specific conditions

b. Sorting the result set in ascending order

c. Performing aggregate calculations on grouped data

d. Joining multiple tables together

Pop Quiz

Q. The SELECT query combined with the GROUP BY clause in SQL is primarily used for _____.

a. Filtering rows based on specific conditions

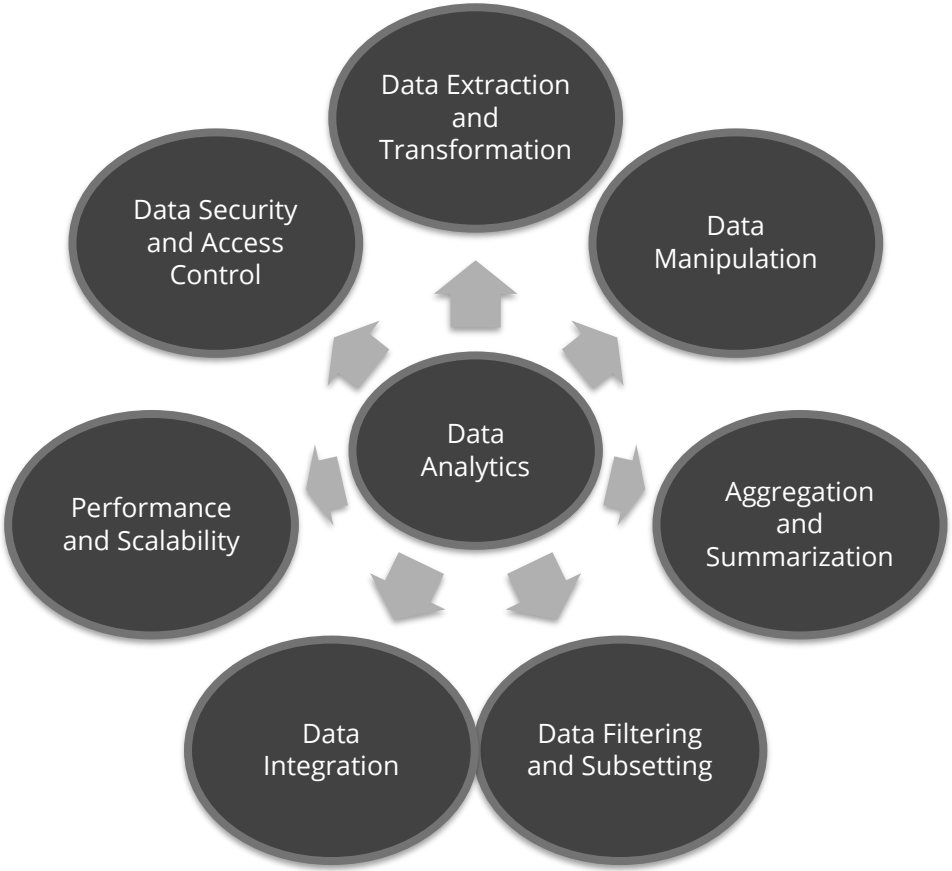
b. Sorting the result set in ascending order

c. Performing aggregate calculations on grouped data

d. Joining multiple tables together

Significance of SQL in Data Analytics

Importance of SQL in Data Analytics



Miscellaneous Commands in SQL

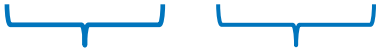
CAST	Used to convert an expression from one data type to another.
DATE	Used to extract the date part from a given date or datetime expression.
CONCAT	Used to concatenate (join together) two or more strings into a single string.
REPLACE	Used to search and replace occurrences of a substring within a string. It replaces all instances of the specified substring with a new substring.

CAST Function in SQL

Purpose and Syntax of CAST Function

Syntax of CAST

```
CAST(expression AS data_type)
```



The value or expression to be converted.

The desired data type to which the expression should be casted.

Example: Casting a numeric value to a different data type:

```
SELECT CAST(25.75 AS
INTEGER);
-- Output: 25

SELECT CAST(10 AS
VARCHAR(10)); -- Output: '10'
```

Demo – CAST Function in SQL

Pop Quiz

Q. Which of the following statements about the CAST function in MySQL is true?

a. The CAST function is used to convert a string to a date or datetime data type

b. The CAST function is used to concatenate multiple strings together

c. The CAST function is used to count the number of rows in a table

d. The CAST function is used to round a numeric value to the nearest whole number

Pop Quiz

Q. Which of the following statements about the CAST function in MySQL is true?

a. The CAST function is used to convert a string to a date or datetime data type

b. The CAST function is used to concatenate multiple strings together

c. The CAST function is used to count the number of rows in a table

d. The CAST function is used to round a numeric value to the nearest whole number

Date and Time Data Type

Purpose and Syntax of Date and Time Data Type

Syntax of DATE

```
DATE(date_expression)
```

The date or datetime expression from which you want to extract the date part.

Example: Extracting the date from a datetime value

```
SELECT DATE('2023-06-22 15:30:00');
-- Output: 2023-06-22
```

Demo – Date and Time Data Type

Pop Quiz

Q. True or False: The DATE function in MySQL can be used to extract the time portion from a given date-time expression.

a. True

b. False

Pop Quiz

Q. True or False: The DATE function in MySQL can be used to extract the time portion from a given date-time expression.

a. True

b. False

CONCAT Function in SQL

Purpose and Syntax of CONCAT Function in SQL

Syntax of DATE

```
CONCAT(string1, string2, ...)
```

The strings or string expressions that you want to concatenate together.

Example: Extracting the date from a datetime value

```
SELECT CONCAT('Hello', 'World');
-- Output: HelloWorld
```

Demo – CONCAT Function in SQL

Pop Quiz

Q. Which of the following options demonstrates the correct usage of the CONCAT function in MySQL to concatenate two strings?

a. CONCAT('Hello', 'World')

b. CONCAT('Hello' + 'World’)

c. CONCATENATE('Hello', 'World’)

d. COMBINE('Hello', 'World')

Pop Quiz

Q. Which of the following options demonstrates the correct usage of the CONCAT function in MySQL to concatenate two strings?

a. **CONCAT('Hello', 'World')**

b. CONCAT('Hello' + 'World')

c. CONCATENATE('Hello', 'World')

d. COMBINE('Hello', 'World')


Q&A

REPLACE Function in SQL

Purpose and Syntax of REPLACE Function in SQL

Syntax of REPLACE

```
REPLACE(string, search_string, replacement_string)
```

The original string in which you want to perform the replacement.

The substring you want to search for within the original string.

The substring that will replace the occurrences of the search string.

Example: To replace a substring

```
SELECT REPLACE('Hello World', 'World', 'Universe'); --
Output: Hello Universe
```

Demo – REPLACE Function in SQL

Pop Quiz

Q. Which of the following options correctly demonstrates the usage of the REPLACE function in MySQL to replace all occurrences of the substring 'abc' with 'xyz' within the string 'abcabcabc'?

a. REPLACE('abcabcabc', 'abc', 'xyz')

b. REPLACE('xyz', 'abc', 'abcabcabc')

c. REPLACE('xyz', 'abcabcabc', 'abc')

d. REPLACE('abc', 'xyz', 'abcabcabc')

Pop Quiz

Q. Which of the following options correctly demonstrates the usage of the REPLACE function in MySQL to replace all occurrences of the substring 'abc' with 'xyz' within the string 'abcabcabc'?

a. **REPLACE('abcabcabc', 'abc', 'xyz')**

b. REPLACE('xyz', 'abc', 'abcabcabc')

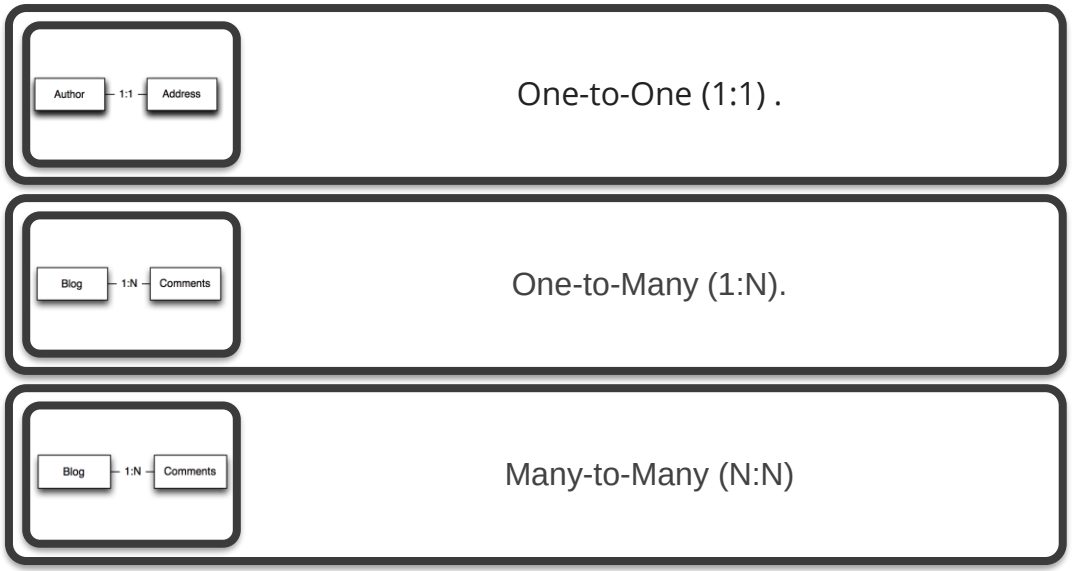
c. REPLACE('xyz', 'abcabcabc', 'abc')

d. REPLACE('abc', 'xyz', 'abcabcabc')

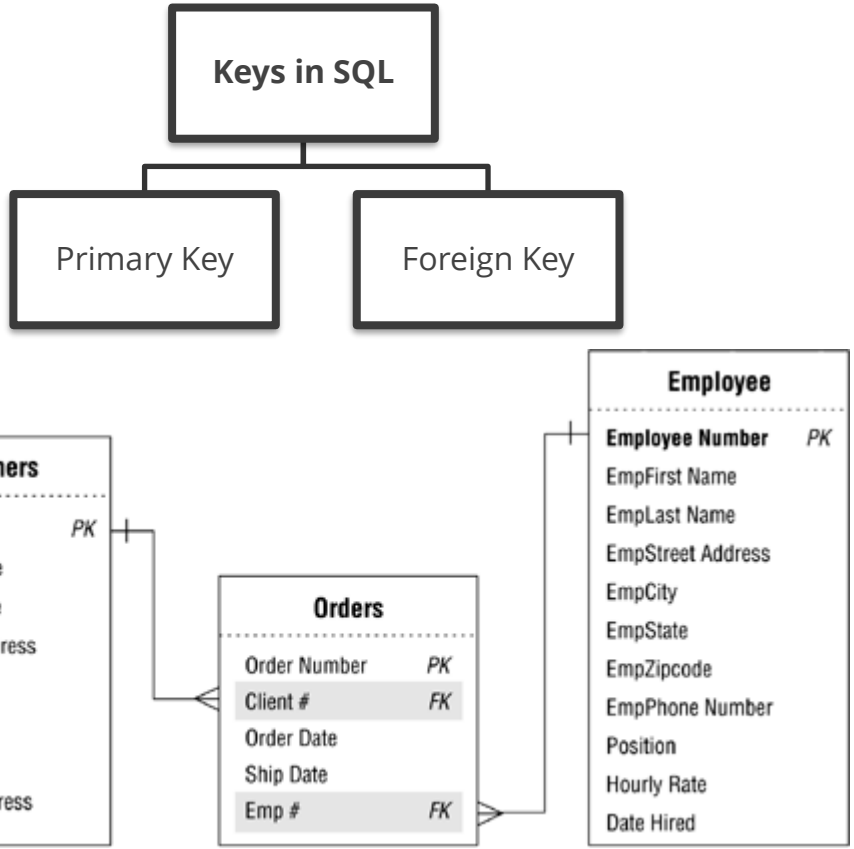
Introduction to Table-to-Table Relationship in SQL

What is Table-to-Table Relationships?

Relationships refer to the associations between tables in a relational database. These relationships define how the data in different tables are related to each other. There are three types of relationships commonly used in MySQL:



Keys in SQL


Keys in SQL
Primary Key
Foreign Key
Customers
Customer ID PK
CustFirst Name
CustLast Name
CustStreet Address
CustState
CustZipcode
CustPhone
CustEmail Address
Orders
Order Number PK
Client # FK
Order Date
Ship Date
Emp # FK
Employee
Employee Number PK
EmpFirst Name
EmpLast Name
EmpStreet Address
EmpCity
EmpState
EmpZipcode
EmpPhone Number
Position
Hourly Rate
Date Hired

Relationships in SQL

Syntax 1 of PRIMARY KEY

```
CREATE TABLE table_name (
column1 datatype,
column2 datatype,
...
PRIMARY KEY (column1)
);
```

Syntax 1 of FOREIGN KEY

```
CREATE TABLE table_name (
column1 datatype,
column2 datatype,
...
FOREIGN KEY (column1) REFERENCES referenced_table (referenced_column) );
```

Syntax 2 of PRIMARY KEY

```
ALTER TABLE table_name ADD
PRIMARY KEY (column1);
```

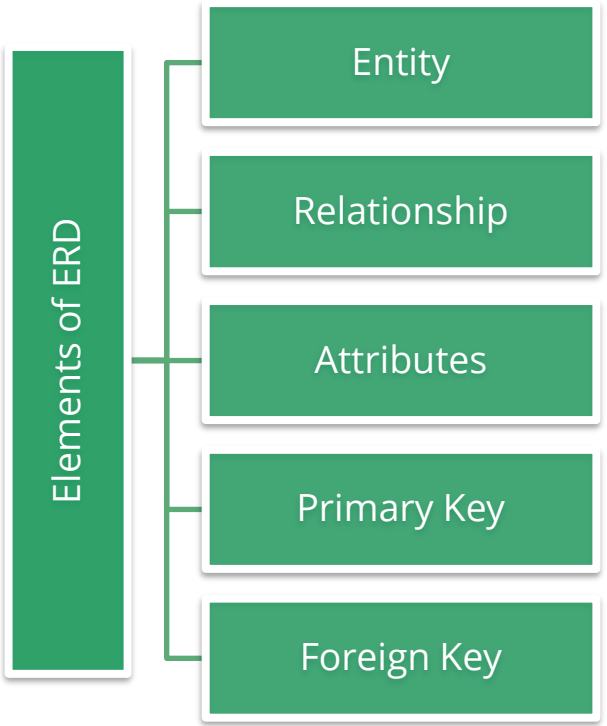
Syntax 2 of FOREIGN KEY

```
ALTER TABLE table_name
ADD FOREIGN KEY (column1) REFERENCES referenced_table (referenced_column);
```

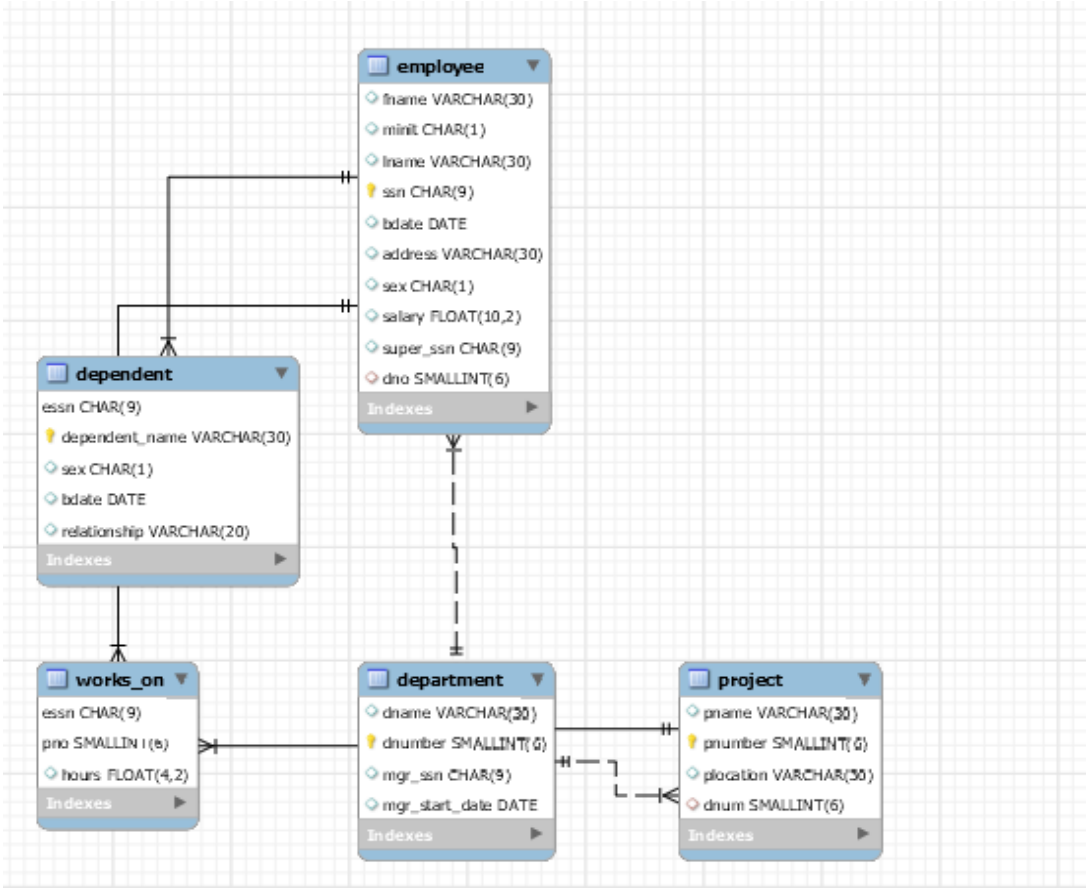
Entity Relationship Diagram

Introduction to ERD

An Entity-Relationship Diagram (ERD) is a visual representation of the entities (objects or concepts) and their relationships within a database. It is a powerful tool used in database design to model and understand the structure of a database system.



Introduction to ERD



Pop Quiz

Q. Which of the following statements about an Entity-Relationship Diagram (ERD) is true?

a. An ERD is used to represent the physical implementation of a database system

b. An ERD is a textual representation of database tables and their relationships

c. An ERD depicts the flow of data within a database system

d. An ERD represents the logical structure of a database system

Pop Quiz

Q. Which of the following statements about an Entity-Relationship Diagram (ERD) is true?

a. An ERD is used to represent the physical implementation of a database system

b. An ERD is a textual representation of database tables and their relationships

c. An ERD depicts the flow of data within a database system

d. An ERD represents the logical structure of a database system


Q&A

Summary

- Recognized the significance of MySQL in data analytics.
- Explored miscellaneous functions in MySQL for data manipulation: CAST, DATE, CONCAT, REPLACE.
- Manipulated text data using CONCAT and REPLACE.
- Gained knowledge about PRIMARY and FOREIGN keys.
- Entity Relationship Diagrams show the blueprint of a database.

Activity 1

Pre-requisites:

- MySQL Workbench

Scenario:

While working on creating a company database, perform the below operations on the existing database.

- Insert a new column **years_of_emp** in the **employee** table which contains the number of years each employee has worked in the organization.
- Set the **employee_id** as the **primary key** of the **employee** table.
- Create a new table called the projects table as shown below. Use **proj_name** as the **primary key**.

proj_name	proj_no	proj_loc
ProductX	1	Bellaire
ProductY	2	Sugarland
ProductZ	3	Houston

Activity 2

Pre-requisites:

- MySQL Workbench

Scenario:

Continue on the previous activity and perform the below operations.

- Create a third table called **works_on** which contains the number of hours each employee has worked on their respective projects.
- Use **emp_id** as the **foreign key** which points to **emp_id** in the **employee** table.
- Use **proj_no** as the **foreign key** which points to **proj_no** in the **projects** table.

emp_id	proj_no	hours
123456789	1	32.5
123456789	2	7.5
333445555	2	10
333445555	3	10
453453453	1	20
453453453	2	20
666884444	3	40

Next Session:
Advanced Data Analysis using SQL

THANK YOU

Please complete your assessments and review the self-learning content for this session on the **PRISM** portal.

knowledgehut
upGrad

knowledgehut
upGrad

Advanced Data Analysis using SQL – Part 1

By the End of this Session, You Will:

- Combine columns from multiple tables using the JOIN clause.
- Concatenate query results vertically using the UNION clause.
- Perform conditional logic within a query using the CASE WHEN statements.
- Perform calculations across a set of rows within a defined window using the WINDOW functions.

Recap

Pop Quiz

Q. Which of the following statements about primary and foreign keys is true?

a. Primary keys are used to establish relationships between tables, while foreign keys uniquely identify records within a table

b. Primary keys and foreign keys are both used to uniquely identify records within a table

c. Primary keys and foreign keys are both used to establish relationships between tables

d. Primary keys uniquely identify records within a table, while foreign keys establish relationships between tables

Pop Quiz

Q. Which of the following statements about primary and foreign keys is true?

a. Primary keys are used to establish relationships between tables, while foreign keys uniquely identify records within a table

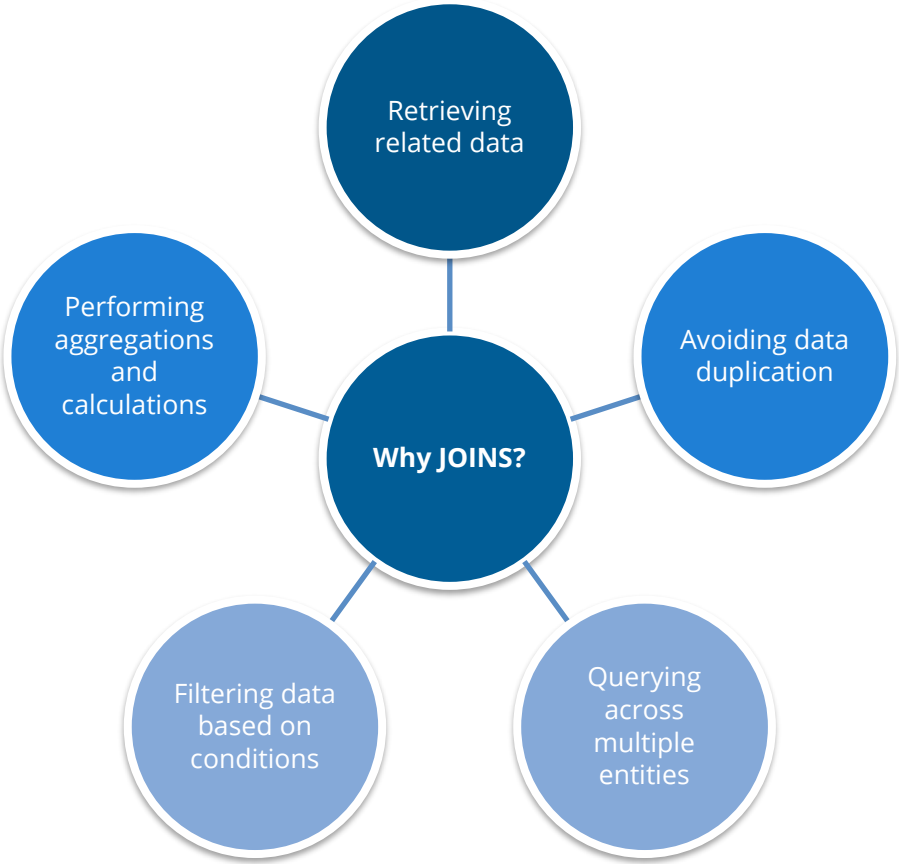
b. Primary keys and foreign keys are both used to uniquely identify records within a table

c. Primary keys and foreign keys are both used to establish relationships between tables

d. Primary keys uniquely identify records within a table, while foreign keys establish relationships between tables

Introduction to Combining Tables

Purpose of Combining Tables



Pop Quiz

Q. Which of the following best describes the purpose of the JOIN clause in SQL?

a. To insert new records into a table

b. To update existing records in a table

c. To retrieve data from multiple tables based on specified relationships

d. To delete records from a table

Pop Quiz

Q. Which of the following best describes the purpose of the JOIN clause in SQL?

a. To insert new records into a table

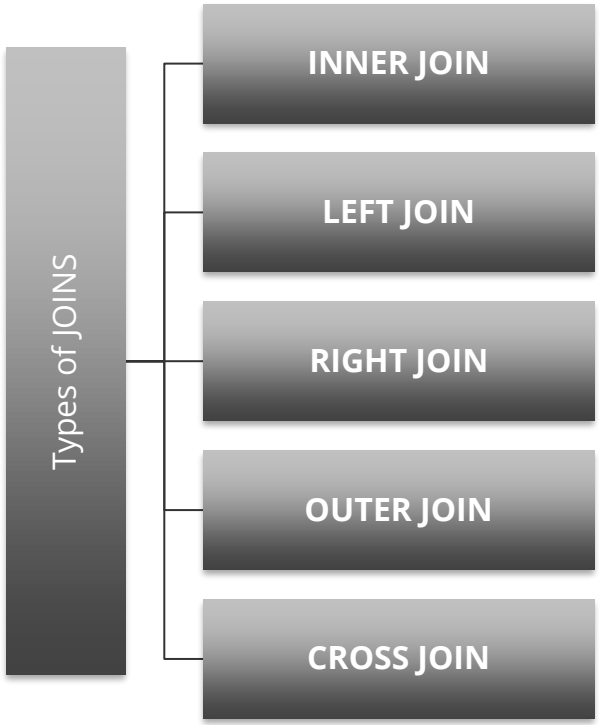
b. To update existing records in a table

c. To retrieve data from multiple tables based on specified relationships

d. To delete records from a table

Introduction to JOIN Clause in SQL

Purpose and Syntax of JOIN Clause



Syntax of JOINS

```
SELECT column(s)
FROM table1
JOIN_TYPE table2
ON table1.column = table2.column;
```

Demo – The JOIN Clause

Pop Quiz

Q. Which of the following JOIN types returns only the matching rows from both tables involved in the JOIN?

a. INNER JOIN

b. LEFT JOIN

c. RIGHT JOIN

d. FULL JOIN

Pop Quiz

Q. Which of the following JOIN types returns only the matching rows from both tables involved in the JOIN?

a. **INNER JOIN**

b. LEFT JOIN

c. RIGHT JOIN

d. FULL JOIN

Introduction to UNION Clause in SQL

Purpose and Syntax of UNION Clause in SQL

The UNION clause in MySQL is used to combine the result sets of two or more SELECT statements into a single result set. It allows you to merge rows from different queries, provided that the column data types and order match.

Syntax of UNION

```
SELECT column1, column2 FROM table1
UNION
SELECT column1, column2 FROM table2;
```

Demo – The UNION Clause in SQL

Pop Quiz

Q. Which of the following statements about the UNION clause in MySQL is true?

a. The UNION clause combines result sets and removes duplicate rows

b. The UNION clause combines result sets and retains duplicate rows

c. The UNION clause is used to join tables based on specified conditions

d. The UNION clause is used to perform aggregate calculations on result sets

Pop Quiz

Q. Which of the following statements about the UNION clause in MySQL is true?

a. **The UNION clause combines result sets and removes duplicate rows**

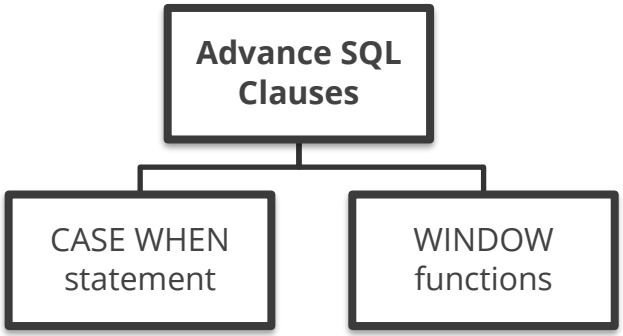
b. The UNION clause combines result sets and retains duplicate rows

c. The UNION clause is used to join tables based on specified conditions

d. The UNION clause is used to perform aggregate calculations on result sets

Introduction to Advanced SQL Data Analysis

Importance of Advanced SQL Techniques




Q&A

CASE Statement in SQL

Purpose and Syntax of CASE Statement in SQL

The CASE WHEN statement in MySQL is a powerful conditional statement that allows you to perform conditional logic within a query.

Syntax of CASE WHEN

```
SELECT column1, column2,
   CASE
      WHEN condition1 THEN result1

      WHEN condition2 THEN result2 ...
      ELSE resultN
   END AS new_column
FROM table_name;
```

Demo – Case Statement in SQL

Pop Quiz

Q. True or False: The CASE WHEN statement in MySQL allows only one condition and one result.

a. False

b. True

Pop Quiz

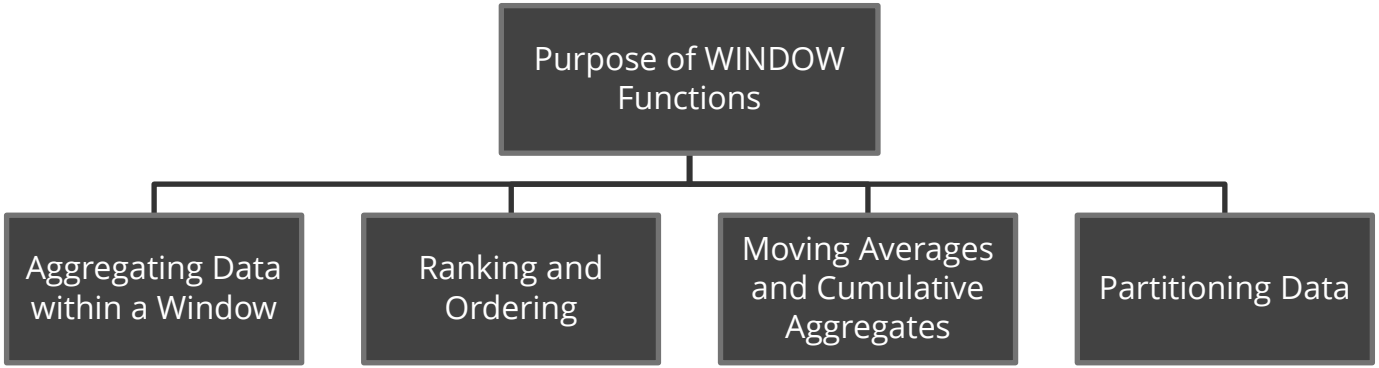
Q. True or False: The CASE WHEN statement in MySQL allows only one condition and one result.

a. False

b. True

Introduction to WINDOW Functions in SQL

Purpose and Syntax of WINDOW Functions in SQL



Syntax of WINDOW Function

```
<window_function> OVER (
   [PARTITION BY <column1>, <column2>, ...]
   [ORDER BY <column1> [ASC|DESC], <column2> [ASC|DESC], ...]
   [ROWS <window_frame>]
)
```

Demo – WINDOW Functions in SQL

Pop Quiz

Q. Which of the following statements about window functions in MySQL is false?

a. Window functions can be used without an ORDER BY clause

b. Window functions can be used with a GROUP BY clause

c. Window functions always require a PARTITION BY clause

d. Window functions can be nested within each other

Pop Quiz

Q. Which of the following statements about window functions in MySQL is false?

a. Window functions can be used without an ORDER BY clause

b. Window functions can be used with a GROUP BY clause

c. Window functions always require a PARTITION BY clause

d. Window functions can be nested within each other


Q&A

Summary

- JOINS in MySQL are of 4 types: INNER, LEFT, RIGHT, and CROSS.
- UNION clause is used to join results from multiple SELECT queries.
- CASE WHEN statements allow us to perform conditional logics within a query.
- WINDOW functions allow us to perform calculations across a set of rows within a defined window.

Activity 1

Pre-requisites:

- MySQL Workbench

Scenario:

Continue on the previous activity and perform the below operations.

- Join the **employee** table with **works_on** table using the **INNER JOIN**.
- While performing the previous join, calculate the average hours spent by employees working in each department.
- Sort the results of the previous query in the descending order of **avg_hours**.
- Filter the departments having **average hours >= 8**.

Session Feedback

Next Session:
Advanced Data Analysis using SQL- Part II

THANK YOU

Please complete your assessments and review the self-learning content for this session on the **PRISM** portal.

knowledgehut
upGrad