

# What is SQL

Structured Query Language, commonly known as SQL, is a robust programming language that is instrumental in managing and manipulating data within relational databases. SQL is a declarative language, used to express and retrieve data efficiently, and in a format that is intuitive and easy to understand. SQL provides a powerful set of commands that can filter, aggregate, and join data from multiple tables and databases, providing valuable insights to data analysts and engineers.

With SQL, you have the ability to interact with data and ask complex questions that can help you understand patterns and trends that might otherwise go unnoticed. SQL is the key to unlocking the full potential of your data, allowing you to analyze it in ways that can have a profound impact on your business.

##Why to learn SQL? Data is the new oil. With the explosion of data in today's world, it has become more important than ever to have the skills to manage and analyze it. SQL is one of the most important tools for working with data, and proficiency in SQL is essential for anyone working in data science.

## 1. Data Retrieval

SQL allows you to extract data from a database quickly and efficiently. With SQL, you can retrieve data that meets specific criteria and filter out data that does not meet those criteria. This makes data retrieval faster and more accurate.

## 2. Data Manipulation

SQL makes it easy to modify data in a database. You can add, delete, or modify data in a database using simple SQL commands. This helps you to clean and transform data in a way that is suitable for analysis.

## 3. Data Analysis

SQL is a powerful tool for data analysis. You can use SQL to aggregate data, calculate statistics, and perform complex queries. This helps you to gain insights into the data and make informed decisions.

## 4. Query Optimization

Data scientists often work with large datasets, and SQL provides powerful tools for querying and analyzing large datasets. By learning SQL, you will be able to optimize your queries to run faster and more efficiently, which is essential for working with big data.

## 5. Real-world Applications

SQL is widely used in the industry for data management and analysis. Almost all businesses and enterprises use databases to store and manage their data. Learning SQL will provide you with the skills to work with databases and manipulate data, which is a valuable skill in any industry.

## 6. Integration with Other Tools

SQL can be integrated with other tools commonly used in data science such as Python, R, and Excel. By learning SQL, you will be able to integrate it with these tools to perform advanced data analysis and build powerful data-driven applications

## How SQL is used in Real World?

Here are some examples of how SQL is used in real-world scenarios:

### E-commerce

Let's say you're working for an e-commerce company that uses a database to store customer data, purchase history, and product information. SQL would be used to query the database and extract insights about customer behavior, popular products, and sales trends.

### Healthcare

In the healthcare industry, SQL is used to manage patient data, clinical trials, and medical research. SQL is used to extract insights about the effectiveness of treatments, patient outcomes, and disease trends.

### Banking

In the banking industry, SQL is used to manage customer accounts, transactions, and credit scores. SQL is used to extract insights about customer behavior, fraud detection, and risk management.

### Marketing

In the marketing industry, SQL is used to manage customer data, campaign performance, and sales data. SQL is used to extract insights about customer demographics, campaign effectiveness, and revenue trends.

In all these industries, SQL is used to extract insights from data and make data-driven decisions. Data scientists who are proficient in SQL are highly sought after by companies of all sizes.

```
# This is formatted as code
```

## Database

A database is anything that collects and organizes data. A spreadsheet holding customer bookings is a database, a plain-text file containing flight schedule data- all can act like a databases. Typically, a database consists of tables, which are made up of rows and columns of data.

Professionally, however, when one refers to a “database” they likely are referring to a database management system (DBMS)

Hence, we can say - A database is a collection of data that is organized in a specific way so that it can be easily accessed, managed, and updated.

And DBMS or Database Management Systems are used to store and manage large amounts of databases for various purposes such as inventory management, financial transactions, customer information, and scientific research.

## RDBMS - Relational Database Management Systems

A relational database management system (RDBMS) is a type of software used to manage relational databases. An RDBMS stores data in the form of tables, which are related to each other through key fields. These key fields are used to create relationships between the tables and enable the RDBMS to perform complex queries and data manipulations.

Table 1: Actors

This table contains information about actors in Bollywood, including their ID, name, gender, and age. The ActorID column serves as the primary key for this table.

ActorID	Name	Gender	Age
1	Shah Rukh Khan	Male	56
2	Deepika Padukone	Female	36
3	Ranbir Kapoor	Male	39

Table 2: Movies

This table contains information about movies in Bollywood, including their ID, title, year of release, and rating. The MovieID column serves as the primary key for this table.

MovieID	Title	Year	Rating
1	DDLJ	1995	8.2
2	Pathaan	2023	8.4
3	Jagga Jasoos	2017	6.4

Table 3: Actor with Movies

This table establishes a many-to-many relationship between actors and movies. Each row represents an actor-movie pairing, and the ActorID and MovieID columns serve as foreign keys that link to the Actors and Movies tables, respectively.

ActorID	MovieID
1	1
1	2

ActorID	MovieID
2	2
3	3

In this example, we can see how the tables are related. The Actors and Movies tables are related to each other through the Actor Movies table, which serves as a bridge table between them. By establishing these relationships, we can perform complex queries and analysis, such as finding all the movies that Shah Rukh Khan has acted in, or finding all the actors who have appeared in a movie released in a particular year.

Overall, tables are the building blocks of a relational database, and they are linked together through key fields to create relationships between them. This enables the RDBMS to perform complex data manipulations and queries, making it a powerful tool for managing and analyzing large datasets.

## Why Separate Tables?

Suppose we have a single table called "Bollywood," which contains information about actors and movies:

ActorID	Name	Gender	Age	Movie ID	Title	Year	Rating
1	Shah Rukh Khan	Male	56	1	DDLJ	1995	8.2
3	Ranbir Kapoor	Male	39	2	Pathaan	2023	8.4
1	Shah Rukh Khan	Male	56	2	Pathaan	2023	8.4
2	Deepika Padukone	Female	36	2	Pathaan	2023	8.4
3	Ranbir Kapoor	Male	39	3	Jagga Jasoos	2017	6.4

This table contains information about actors and movies, but it suffers from several problems. First, it is redundant, meaning that the same actor and movie information is repeated multiple times. This can lead to bloated table sizes and make it difficult to maintain the data.

Second, it violates the principle of normalization, which is the process of organizing data into tables to eliminate redundancy and improve data integrity. By separating the data into multiple tables, we can eliminate redundancy and ensure that each piece of information is stored in only one place.

## Why MySQL?

MySQL is an excellent choice for learning SQL for several reasons:

## Popularity

MySQL is one of the most widely used relational database management systems (RDBMS) in the world. It is used by many organizations, from small startups to large enterprises, which means that there are many job opportunities available for those with MySQL skills.

## Ease of Use

MySQL is known for its ease of use and simplicity. Its syntax is straightforward and easy to understand, making it a great choice for beginners who are just starting to learn SQL.

## Availability

MySQL is available for free, which means that anyone can download and install it on their computer. This makes it an accessible choice for students and self-learners who want to practice SQL on their own.

## Community

MySQL has a large and active community of users and developers who are constantly creating new resources and tools for learning and working with MySQL. This includes online forums, tutorials, and open-source projects.

## Compatibility

MySQL is compatible with a wide range of programming languages, including PHP, Java, Python, and C++. This makes it a versatile choice for developers who want to work with different languages and platforms.

Overall, MySQL is a great choice for learning SQL because of its popularity, ease of use, availability, community, and compatibility. Whether you are a student, a self-learner, or a developer, MySQL provides a solid foundation for learning and working with SQL.

## MySQL Workbench

MySQL Workbench is a visual tool for designing, developing, and administering MySQL databases. It provides a graphical user interface (GUI) for managing MySQL databases, making it easier for users to create and modify database structures, write and execute SQL queries, and perform other database-related tasks.

MySQL Workbench is popular in the MySQL community because it is a comprehensive tool that simplifies the database development and administration process. Some of the key features of MySQL Workbench include:

- **Visual database design and modeling:** MySQL Workbench allows users to create and edit database structures visually, using drag-and-drop tools and an intuitive interface.

- SQL code editor: MySQL Workbench includes a powerful SQL code editor with syntax highlighting, autocomplete, and code snippets, which makes it easy to write and execute SQL queries.
- Query builder: MySQL Workbench provides a query builder that allows users to create and execute SQL queries visually, without having to write any code.
- Database administration tools: MySQL Workbench includes tools for managing user accounts, backups, and performance, which makes it easier for administrators to manage their MySQL databases.
- Database migration tools: MySQL Workbench also includes tools for migrating databases from other systems to MySQL, which simplifies the process of switching to MySQL.

MySQL Workbench comes pre-installed with MySQL server because it is developed by the same company that develops MySQL, Oracle Corporation. This integration ensures that MySQL Workbench is always up-to-date and fully compatible with MySQL server.

## Installing MySQL Workbench

Open the MySQL website on a browser. Click on the following link -

<https://dev.mysql.com/downloads/windows/installer/8.0.html>

Choose your operating system and follow installation guidelines.

*No need to create an account. Just click on the link "No thanks, just start my download."*

## Dummy Database to Practise

- Open MySQL Workbench and connect to your MySQL server.
- Create a new query tab by clicking on the "New Query" button in the toolbar.
- Copy and paste the SQL code into the new query tab.
- Click on the "Execute" button in the toolbar to run the SQL code.
- If the code executes successfully, you should see a message indicating that the queries were executed without any errors.
- You can then check the tables and data by opening the "Schema" tab in the sidebar, selecting your database, and expanding the tables to view their columns and data.

```
-- Drop the database if it already exists
DROP DATABASE IF EXISTS bollywood_movies;

-- Create the database
CREATE DATABASE bollywood_movies;
```

```

-- Use the database
USE bollywood_movies;

-- Create the movies table
CREATE TABLE movies (
    movie_id INT NOT NULL AUTO_INCREMENT,
    title VARCHAR(255) NOT NULL,
    year INT NOT NULL,
    genre VARCHAR(255) NOT NULL,
    director VARCHAR(255) NOT NULL,
    PRIMARY KEY (movie_id)
);

-- Create the actors table
CREATE TABLE actors (
    actor_id INT NOT NULL AUTO_INCREMENT,
    name VARCHAR(255) NOT NULL,
    gender ENUM('Male', 'Female') NOT NULL,
    age INT NOT NULL,
    PRIMARY KEY (actor_id)
);

-- Create the movie_actors table to establish a many-to-many
relationship
CREATE TABLE movie_actors (
    movie_id INT NOT NULL,
    actor_id INT NOT NULL
);

-- Insert data into the movies table
INSERT INTO movies (title, year, genre, director)
VALUES
('Dilwale Dulhania Le Jayenge', 1995, 'Romance', 'Aditya Chopra'),
('Kuch Kuch Hota Hai', 1998, 'Romance', 'Karan Johar'),
('Hum Aapke Hain Koun..!', 1994, 'Romance', 'Sooraj Barjatya'),
('Darr', 1993, 'Thriller', 'Yash Chopra'),
('Khal Nayak', 1993, 'Action', 'Subhash Ghai');

-- Insert data into the actors table
INSERT INTO actors (name, gender, age)
VALUES
('Shah Rukh Khan', 'Male', 56),
('Kajol', 'Female', 47),
('Madhuri Dixit', 'Female', 54),
('Salman Khan', 'Male', 55),
('Sanjay Dutt', 'Male', 62);

-- Insert data into the movie_actors table
INSERT INTO movie_actors (movie_id, actor_id)
VALUES

```

```
(1, 1),  
(1, 2),  
(2, 1),  
(2, 2),  
(3, 4),  
(3, 3),  
(4, 1),  
(4, 5),  
(5, 4),  
(5, 1);
```

## ## Test Yourself

### What does SQL stand for?

- a. Standard Query Language
- b. Structured Query Language
- - c. Simple Query Language

d. System Query Language Answer: b. Structured Query Language

Explanation: SQL stands for Structured Query Language, which is a programming language used for managing and manipulating data in relational databases.

### Why is it important to learn SQL?

- a. To improve typing speed
- b. To communicate with people
- c. To manage and analyze data
- d. To design websites Answer: c. To manage and analyze data

Explanation: With the explosion of data in today's world, it has become more important than ever to have the skills to manage and analyze it. SQL is one of the most important tools for working with data, and proficiency in SQL is essential for anyone working in data science.

### What is the primary function of SQL?

- a. Data visualization
- b. Data manipulation
- c. Data storage
- d. Data processing Answer: b. Data manipulation

Explanation: SQL makes it easy to modify data in a database. You can add, delete, or modify data in a database using simple SQL commands. This helps you to clean and transform data in a way that is suitable for analysis.

### What is a database?

- a. A collection of data that is organized in a specific way
- b. A programming language used for managing data



- c. A tool for data visualization
- d. A method for data storage Answer: a. A collection of data that is organized in a specific way

Explanation: A database is a collection of data that is organized in a specific way so that it can be easily accessed, managed, and updated.

## What is an RDBMS?

- a. A tool for database design and modeling
- b. A type of software used to manage relational databases
- c. A method for data visualization
- d. A programming language used for managing data Answer: b. A type of software used to manage relational databases

Explanation: A relational database management system (RDBMS) is a type of software used to manage relational databases. An RDBMS stores data in the form of tables, which are related to each other through key fields.

## What is the primary benefit of separating data into multiple tables?

- a. Faster data retrieval
- b. Reduced redundancy
- c. More accurate data
- d. All of the above Answer: d. All of the above

Explanation: By separating the data into multiple tables, we can eliminate redundancy and ensure that each piece of information is stored in only one place. This makes data retrieval faster and more accurate.

## What is the primary benefit of using MySQL for learning SQL?

- a. Popularity
- b. Ease of use
- c. Availability
- d. All of the above Answer: d. All of the above

Explanation: MySQL is one of the most widely used relational database management systems (RDBMS) in the world, and it is known for its ease of use and simplicity. It is available for free, and it is compatible with a wide range of programming languages.

## What is MySQL Workbench?

- a. A programming language used for managing data
- b. A visual tool for designing, developing, and administering MySQL databases
- c. A tool for data visualization
- d. A method for data storage Answer: b. A visual tool for designing, developing, and administering MySQL databases

Explanation: MySQL Workbench is a visual tool for designing, developing, and administering MySQL databases. It provides a graphical user interface (GUI) for managing MySQL databases,

making it easier for users to create and modify database structures, write and execute SQL queries, and perform other database-related tasks.

### What is the purpose of using SQL in data analysis?

- A) To store and manage data in a database
- B) To retrieve data from a database
- C) To modify data in a database
- D) To perform complex queries and calculations on data

Answer: D

Explanation: SQL is a powerful tool for data analysis as it allows for the manipulation of data through complex queries and calculations. It can help extract insights from large datasets and make data-driven decisions.

### Which of the following industries use SQL to manage their data?

- A) E-commerce
- B) Healthcare
- C) Banking
- D) Marketing
- E) All of the above

Answer: E

Explanation: SQL is used in a wide range of industries for managing and analyzing data. E-commerce companies use it to extract insights about customer behavior and sales trends, healthcare companies use it to manage patient data and research, banks use it to manage customer accounts and detect fraud, and marketing companies use it to analyze customer data and campaign performance.

