

# SQL Vs NoSQL

## 1.What is SQL?

SQL, which stands for Structured Query Language, is a domain-specific programming language (e.g., a language targeted to a specific task or problem) that is commonly used for tasks such as inserting, updating, querying, and deleting data within a database. SQL is also used to create and modify database schemas (e.g., data formatting rules, table/index structure ) as well as define database access and administration parameters.

- **Structured Data:** Data is organized in tables with rows and columns, making it easy to relate different types of information.
- **ACID Compliance:** SQL databases follow the [ACID](#) properties (Atomicity, Consistency, Isolation, Durability) to ensure reliable transactions and data integrity.
- **Examples:** Popular SQL databases include **MySQL**, **PostgreSQL**, **Oracle**, and **MS SQL Server**.

### 1.1What is structured data?

Structured data is data that is organized in a consistent, predefined format and often consists of alphanumeric characters. Examples include financial transactions, inventory records, or customer lists which are often stored in SQL databases (e.g., relational databases).

## 2. What is NoSQL?

**NoSQL**, which stands for Not only SQL, is a database management system approach used to ingest, store, and retrieve unstructured data and semi-structured data within a database. This means that data that cannot be analyzed or counted through traditional relational databases (e.g., SQL) can remain in its native format and be ingested into a NoSQL database.

The reason it is called NoSQL is to emphasize that these databases can handle non-tabular, non-relational data models as well as support SQL-like query languages.

- **Flexible Schema:** NoSQL [databases](#) allow the storage of data without a predefined structure, making them more adaptable to changing data requirements.
- **CAP Theorem:** NoSQL databases are designed based on the **CAP theorem** (Consistency, Availability, Partition Tolerance), which prioritizes availability and partition tolerance over strict consistency.
- **Examples:** Well-known NoSQL databases include **MongoDB**, **Cassandra**, **CouchDB**, and **HBase**.

## 2.2 What is unstructured data?

Unstructured data is data that doesn't have a predefined data model or consistent organization. In addition, unstructured data, such as social media posts, can update and change rapidly while structured data, such as bank transactions, have a much lower rate of change. Examples of unstructured data include pictures, audio files, videos, and maps.

### SQL vs NoSQL:

- **SQL Databases:** Generally, SQL databases perform well for **complex queries**, structured data, and systems requiring **data consistency** and **integrity**. However, as the volume of data grows, they may struggle with **scalability** and may require significant infrastructure upgrades.
- **NoSQL Databases:** NoSQL databases excel in scenarios that demand **high performance** and **scalability**. Because of their **horizontal scalability** (accommodating more servers), they handle large amounts of data and high-velocity workloads better. For instance, MongoDB or Cassandra is a common choice when dealing with [big data](#) or applications with high traffic.

### 3. Key Features of MongoDB

#### 1. Document-Oriented Storage:

- MongoDB stores data in the form of **documents** using **BSON** (Binary JSON).
- Each document is a key-value pair, allowing complex nested data (like arrays and sub-documents).

#### 2. Flexible Schema:

- No need to define the structure of documents in advance.
- Different documents in the same collection can have different fields and data types.
- This allows for agile and iterative development.

#### 3. Horizontal Scalability with Sharding:

- Supports **sharding** (partitioning data across multiple servers).
- Automatically balances data across shards.
- Enables scaling out to handle large volumes of data and traffic.

#### 4. MongoDB Atlas (Cloud Database):

- Managed cloud database service provided by MongoDB.
- Offers:
  - Automated backups
  - Real-time monitoring
  - Auto-scaling
  - Global clusters
  - Security integrations

