**What is NoSQL Database**

* **Why NoSQL**

1. NoSQL is a database type that stores information(data) in JSON documents format instead of tables(columns and rows) like relational databases.
2. NoSQL databases are more flexible, scalable. Used to deal with large(complex) data problems.
3. It does not require a schema and used for big data and real-time web applications.
4. The main purpose to build NoSQL database is to use the large data storage needs over distributed data stores.
5. These databases are faster than SQL(relational databases) .

* **History about NoSQL Databases**

1. The term NoSQL was first used in 1998 by Carlo Strozzi while naming his lightweight, open-source relational database which does not use SQL.
2. The term NoSQL developed as a response to web data, the need for processing unstructured data, and for faster processing. It uses distributed database system, meaning a data over multiple computers.
3. In research, with finding that NoSQL database is a better choice for large, unstructured data, and provides speed and flexibility compared to relational databases.
4. This led to organizations such as Facebook,Twitter, LinkedIn, and Google adopting NoSQL systems. These organizations need large amount of unstructured data to process, Big data became an official term in 2005.
5. In 2009, the term NoSQL was introduced again by Eric Evans and Johan Oskarsson.

* **Features of NoSQL**

Some of the features are mentioned below:

1. **Multi-Model**

* As relational databases need data to stored in multiple tables to be accessed and analysed, but NoSQL database model makes it more flexible to handle the data.
* NoSQL database can deal with structured as well as unstructured data with the same ease. NoSQL database store the data over multiple servers, which allows to use the same data in different data models without creating any complexity.

1. **Easily Scalable**

* The Sharding process is about dividing the database into smaller parts(chunks) over multiple servers, and this leads to easy scaling to the large volume data.
* The scalability also improves the performance, high availability of data. It’s not like relational databases cannot scale, but NoSQL databases are more scalable than relational one.

1. **Flexible**

* The NoSQL databases are more flexible as it stores the data as documents and not in multiple tables. It does not require any schema, as it is a unstructured data compared to relational database which requires structured data.

1. **Distributed**

* The data in NoSQL databases are distributed data as it is stored all over multiple servers with multiple locations.
* The main advantage of using a distributed database is that it provides high availability of data by providing multiple copies of data.
* **What is the CAP Theorem?**

CAP theorem is also known as Brewer’s theorem named after computer scientist Eric Brewer.

CAP stands for Consistency, Availability and Partition Tolerance. It states that it is not possible for a distributed data systems to simultaneously provide more than two out of these three guarantees.

* **Consistency**: Consistency is all about the consistent data. The data should be remain consistent even after the execution or any update in data. Even after updating the data, all clients should be able to see the same data.
* **Availability**: The data from any database should always be available and be responsive.
* **Partition Tolerance**: It means that the system should continue to function in case there is no stability among the communication between the servers. The servers can be partitioned into multiple groups which may not communicate with each other. And if some part of the database is unavailable, the rest of the parts will not be affected.
* **Eventual Consistency**

This means to have copies of data on multiple machines from which we get high availability and scalability. Hence if changes made to any machine’s data, this also has to make changes to other data.

It follows the rule called BASE, which is :

Basically Available, Soft state, Eventual consistency.

* Basically available means the database has to be available all the time as per the CAP theorem.
* Soft state means even if there is no input given, the system state may change.
* Eventual consistency means that the system has to be consistent over time.
* **Advanatges of NoSQL**

1. It handles large volume of data at high speed.
2. It is a schema less database
3. Has dynamic schema.
4. Requires less management.
5. Flexible data models
6. Easy replication
7. Provides high performance and easy scaling.
8. It can handle structured and unstructured data as well.
9. Much more flexible.

* **Disadvantages of NoSQL**

1. Less mature compared to RDBMS databases and tools.
2. Less support
3. No standardization rules.
4. Does not support for the consistency for multiple transactions.
5. It is difficult to maintain unique values when there is a increase in data volume.
6. Learning is not that easy for new developers.