1. **What is NoSQL data base?**

NoSQL is an approach to databases that represents a shift away from traditional relational database management systems (RDBMS). Relational databases rely on tables, columns, rows, or schemas to organize and retrieve data. In contrast, NoSQL databases do not rely on these structures and use more flexible data models. NoSQL can mean “not SQL” or “not only SQL.” As RDBMS have increasingly failed to meet the performance, scalability, and flexibility needs that next-generation, data-intensive applications require, NoSQL databases have been adopted by mainstream enterprises. NoSQL is particularly useful for storing unstructured data, which is growing far more rapidly than structured data and does not fit the relational schemas of RDBMS.

**2. How does data get stored in NoSQL database?**

There are different ways to store data in NoSQL database, one of them is Key-value NoSQL databases which is very useful for an application to support high-speed read and write processing of non-transactional data. Stored values can be any type of binary object (text, video, JSON document, etc.) and are accessed via a key. The application has complete control over what is stored in the value, making this the most flexible NoSQL model. Data is partitioned and replicated across a cluster to get scalability and availability. For this reason, key value stores often do not support transactions. However, they are highly effective at scaling applications that deals with high-velocity, non-transactional data.

**3. What is a column family in HBase?**

In the HBase data model columns are grouped into column families, which must be defined up front during table creation. Column families are stored together on disk, which is why HBase is referred to as a column-oriented data store.

**4. How many maximum number of columns can be added to HBase table?**

There is no hard limit on maximum number but HBase currently does not do well with anything above two or three column families so it is recommended to keep the number of column families in the schema low.

**5. Why columns are not defined at the time of table creation in HBase?**

To maintain the dynamic behavior of columns, columns are not defined at the time of table creation. For Example if we are having address as a column family, we can have different column qualifiers such as city, state etc. at run time.

**6. How does data get managed in HBase?**

Just like in a Relational Database, data in HBase is stored in Tables and these Tables are stored in Regions. When a Table becomes too big, the Table is partitioned into multiple Regions. These Regions are assigned to Region Servers across the cluster. Each Region Server hosts roughly the same number of Regions.

**7. What happens internally when new data gets inserted into HBase table?**

Each Region Server contains a Write-Ahead Log (called HLog) and multiple Regions. Each Region in turn is made up of a MemStore and multiple StoreFiles (HFile). The data lives in these StoreFiles in the form of Column Families. The MemStore holds in-memory modifications to the Store (data).