Task 1 : Questions and Answers

1) NO SQL DATABASE:
--> Also called as "not only sql database", it is required to process bigdata and at the same time it supports the low latency queries.

2) How the data is stored in no sql database
--> The data is stored as :

There are 3 types of no sql database:

1) Columnar no sql database: Here the structured data are stored(like testfiles) in rows and columns. Ex HBASE and CASSENDRA

2) Documental no sql database: Here the semi structured data is stored(like json,xml etc). Ex Mongodb and Mark logic

3) Memory no sql database: The data is stored in memory(Structured data). Ex MEMCACHE,COUCHDB AND REDIS DB.

4) Graph no sql database: The data is stored using graph structures with nodes and edges. Ex NEO4J.

3) What is a column family in HBase?
--> The group of columns or collection columns in table or database.

4) How many maximum number of columns can be added to HBase table?
--> HBASE cannot process well above two or three column families. SO there is should be low column families.

5) How does data get managed in HBase?
-->HBase is a column-oriented database that's an open-source implementation of Google's Big Table storage architecture. It can manage structured and semi-structured data and has some built-in features such as scalability, versioning, compression and garbage collection.