**ASSIGNMENT -31.3**

**Problem Statement: Explain in brief**

1. **Hbase is a schema less database, what does it mean?**

Without HBase you can’t do table joins and that’s why it is considered as “schema-less” database.

HBase is a wide column store and has column families which are roughly equivalent to tables.  The column names can be completely variable and the number of columns can vary by row so you could have a table with billions of rows and could have rows with million columns.

1. **What is the minimum number of column family every Hbase table should have?**

Minimum number of column family every HBASE table should is 1 because HBASE stores the name of the column under the column family only.

Column family make it very simple in finding the particular record. Also for each created data in hbase creates the timestamp, value length and other details so it is necessary to create at least the column family.

Also HBASE has schema less structure and hence for creating a table it is created with only a column family and while inserting the values the column name is mentioned.

1. **What is the benefit of using connection pool in Hbase?**

The connection object contains logic to find the master, locate regions out on the cluster, keeps a cache of locations and then knows how to re-calibrate after they move. The individual connections to servers, meta cache, zookeeper connection, etc. are all shared by the [Table](https://hbase.apache.org/apidocs/org/apache/hadoop/hbase/client/Table.html) and [Admin](https://hbase.apache.org/apidocs/org/apache/hadoop/hbase/client/Admin.html) instances obtained from this connection.

Connection creation is a heavy-weight operation. Connection implementations are thread-safe, so that the client can create a connection once, and share it with different threads. [Table](https://hbase.apache.org/apidocs/org/apache/hadoop/hbase/client/Table.html) and [Admin](https://hbase.apache.org/apidocs/org/apache/hadoop/hbase/client/Admin.html) instances, on the other hand, are light-weight and are not thread-safe.

Typically, a single connection per client application is instantiated and every thread will obtain its own Table instance. Caching or pooling of [Table](https://hbase.apache.org/apidocs/org/apache/hadoop/hbase/client/Table.html) and [Admin](https://hbase.apache.org/apidocs/org/apache/hadoop/hbase/client/Admin.html) is not recommended.