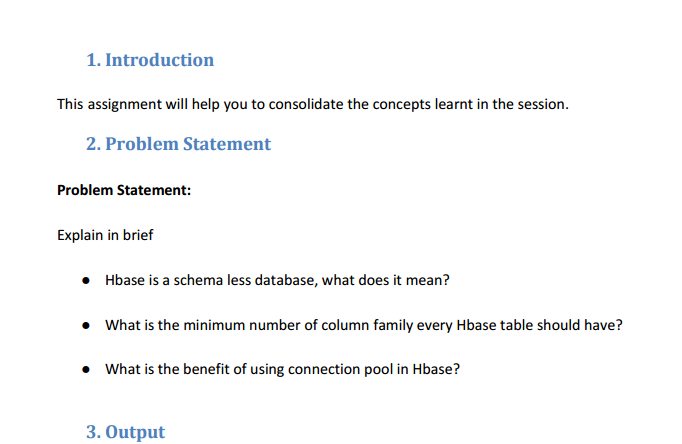
**Assignment 31.3**



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

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

* **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. 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.

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.

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

Opening/Closing Hbase connections is an expensive process and hence connection pools improve the performance of execution of commands on a database for which we maintain connection objects in the pool. It facilitates reuse of the same connection object to serve a number of client requests.

Every time a client request is received, the pool is searched for an available connection object and it's highly likely that it gets a free connection object.

Otherwise, either the incoming requests are queued or a new connection object is created and added to the pool (depending upon how many connections are already there in the pool and how many the particular implementation and configuration can support).

As soon as a request finishes using a connection object, the object is given back to the pool from where it's assigned to one of the queued requests (based on what scheduling algorithm the particular connection pool implementation follows for serving queued requests). Since most of the requests are served using existing connection objects only so the connection pooling approach brings down the average time required for the users to wait for establishing the connection to the database.

Java Code to create Connection In HBase

// Create a connection to the cluster.

HConnection connection = HConnectionManager.createConnection(Configuration);

HTableInterface table = connection.getTable("myTable");

// use table as needed, the table returned is lightweight

table.close();

// use the connection for other access to the cluster

connection.close();