**Lab 4 (MYSQL, Hbase, sqoop)**

**Name: Zeyad Mohamed ID: 21ZTEM**

Step 1: open **hbase** **shell** (I Just opened hbase shell to start inserting **NoSQL** commands).

Graphical user interface, text, application

Description automatically generated

Step 2: Create **‘Customers’** table with two column families **‘CustomerInfo’** and **‘ContactInfo’** , using **create** command**.**

A picture containing chart

Description automatically generated

Step 3: verify that table is already created.

Background pattern

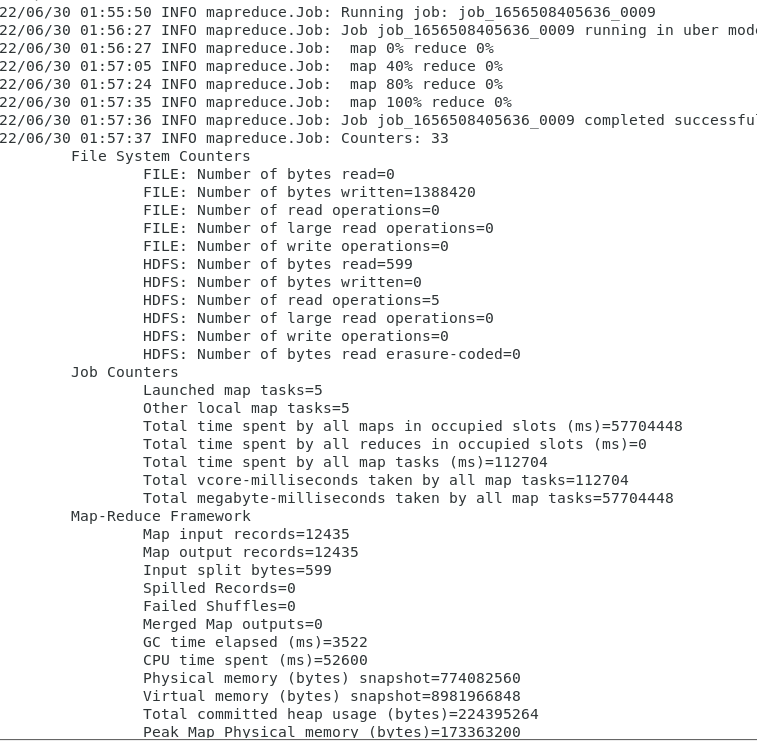
Description automatically generated with low confidence

Step 4: Start import required data and columns from MySQL **retail\_db.customers** table to **Customers** database in hbase, using **sqoop CLI commands.**

1. Put (customer\_fname, customer\_lname, customer\_email, customer\_password ) columns data under **CustomerInfo** column family.

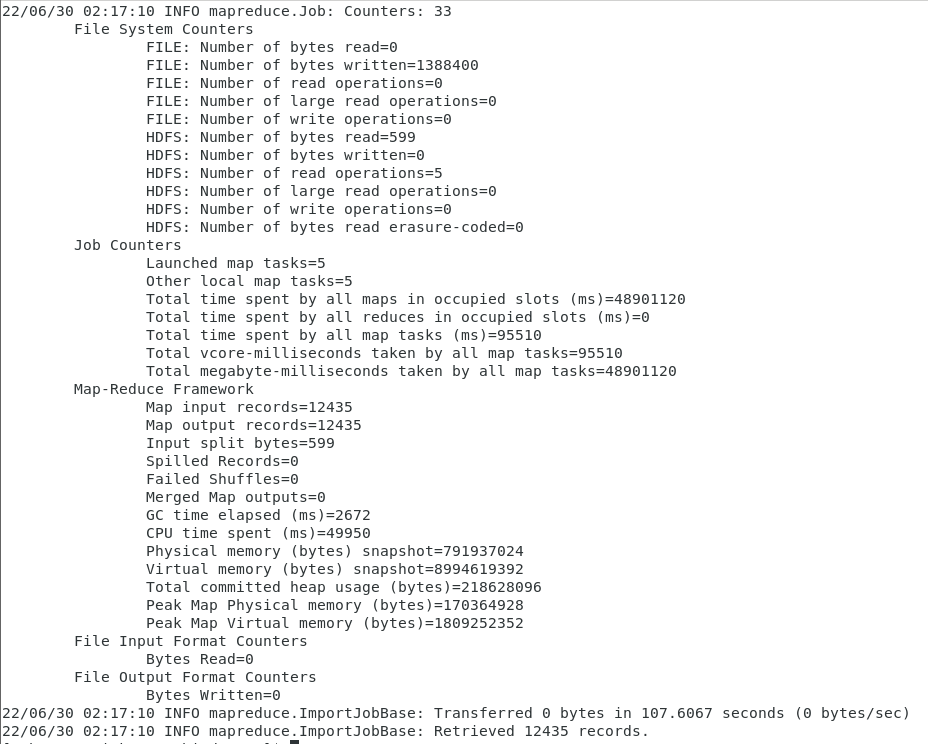


Output:

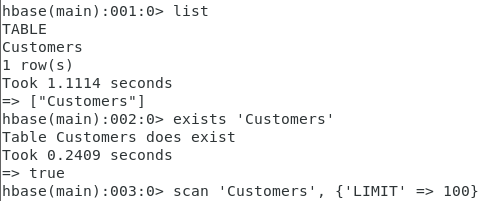


1. put (customer\_street, customer\_city, customer\_state, customer\_zipcode) columns data under **ContactInfo** column family.

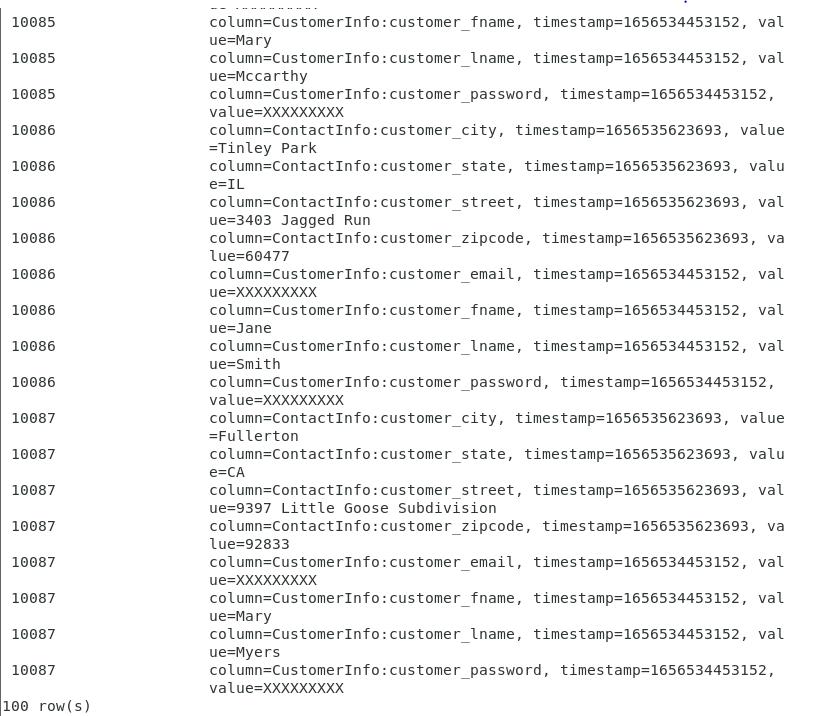




Step 5: verify the existence of the table and the data in the **Customers** dataset using **hbase** commands, to check if the **sqoop** worked well or not.



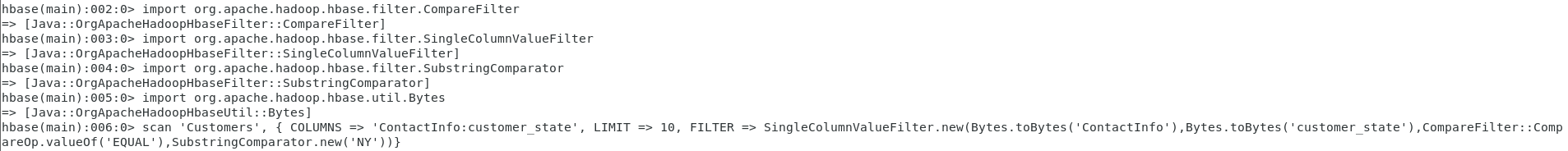
Output:



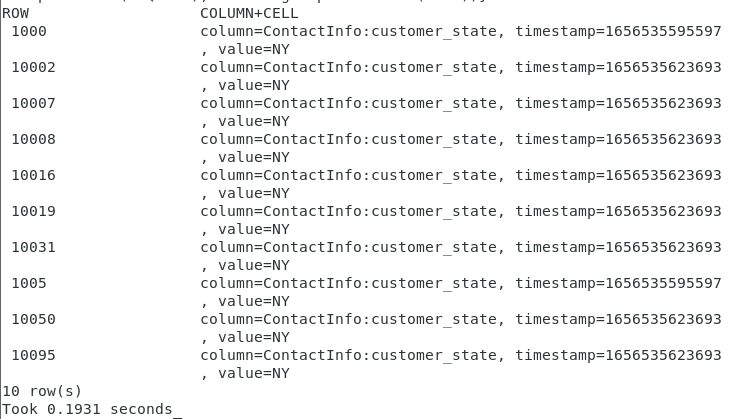
Step 6: List only top 10 records of NY in customers database.

In this and next step we will use four packages in Hbase to solve this problem.

1. CompareFilter: This is a generic filter to be used to filter by comparison. It takes an operator (equal, greater, not equal, etc.) and a byte [] comparator.
2. SingleColumnValueFilter: This filter is used to filter cells based on value. It takes a [CompareOperator](https://hbase.apache.org/apidocs/org/apache/hadoop/hbase/CompareOperator.html) operator (equal, greater, not equal, etc), and either a byte [] value or a ByteArrayComparable.
3. SubstringComparator: This comparator is for use with SingleColumnValueFilter, for filtering based on the value of a given column. Use it to test if a given substring appears in a cell value in the column. The comparison is case insensitive.
4. Bytes: Utility class that handles byte arrays, conversions to/from other types, comparisons, hash code generation, manufacturing keys for HashMaps or HashSets, and can be used as key in maps or trees.



Output:



Step 7: List only top 10 records of VA in customers database.

I’ll will use the same previous command, but I’ll change the value which passed to **Substringoperator.new()**



Output:

Text

Description automatically generated with medium confidence

REF:

1. https://hbase.apache.org/apidocs/org/apache/hadoop/hbase/util