

## Creating Lookup Tables

1. Creating "CC\_lookup\_data\_hbase" table (hive-hbase integrated table) which will be visible in HBase (lookup\_data\_hive).

```
CREATE TABLE CC_LOOKUP_DATA_HBASE(`CARD_ID` STRING,`UCL` DOUBLE, `SCORE` INT,
`POSTCODE`
STRING, `TRANSACTION_DT` TIMESTAMP) STORED BY
'org.apache.hadoop.hive.hbase.HBaseStorageHandler' WITH SERDEPROPERTIES
("hbase.columns.mapping"=":key, lookup_card_family:ucl,
lookup_card_family:score,lookup_transactions_family:postcode,
lookup_transactions_family:transactions_dt") TBLPROPERTIES
("hbase.table.name" = "cc_lookup_data_hive");
```

```
hive> CREATE TABLE CC_LOOKUP_DATA_HBASE(`CARD_ID` STRING,`UCL` DOUBLE, `SCORE` INT, `POSTCODE`
> STRING, `TRANSACTION_DT` TIMESTAMP) STORED BY
> 'org.apache.hadoop.hive.hbase.HBaseStorageHandler' WITH SERDEPROPERTIES
> ("hbase.columns.mapping"=":key, lookup_card_family:ucl, lookup_card_family:score,
> lookup_transactions_family:postcode, lookup_transactions_family:transactions_dt") TBLPROPERTIES
> ("hbase.table.name" = "cc_lookup_data_hive");
OK
Time taken: 1.65 seconds
hive>
```

2. Verifying details of **cc\_lookup\_data\_hive** (hive-hbase integrated) table:

**describe 'cc\_lookup\_data\_hive'**

```
hive> CREATE TABLE CC_LOOKUP_DATA_HBASE(`CARD_ID` STRING,`UCL` DOUBLE, `SCORE` INT, `POSTCODE`
> STRING, `TRANSACTION_DT` TIMESTAMP) STORED BY
> 'org.apache.hadoop.hive.hbase.HBaseStorageHandler' WITH SERDEPROPERTIES
> ("hbase.columns.mapping"=":key, lookup_card_family:ucl, lookup_card_family:score,
> lookup_transactions_family:postcode, lookup_transactions_family:transactions_dt") TBLPROPERTIES
> ("hbase.table.name" = "cc_lookup_data_hive");
OK
Time taken: 1.65 seconds
hive> exit;
[root@ip-172-31-21-198 ec2-user]# hbase shell
HBase Shell
Use "help" to get list of supported commands.
Use "exit" to quit this interactive shell.
Version 1.4.13, rUnknown, Fri Apr 17 15:18:24 UTC 2020

hbase(main):001:0> describe 'cc_lookup_data_hive'
Table cc_lookup_data_hive is ENABLED
cc_lookup_data_hive
COLUMN FAMILIES DESCRIPTION
{NAME => 'lookup_card_family', BLOOMFILTER => 'ROW', VERSIONS => '1', IN_MEMORY => 'false', KEEP_DELETED
_CELLS => 'FALSE', DATA_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER', COMPRESSION => 'NONE', MIN_VERSIONS
=> '0', BLOCKCACHE => 'true', BLOCKSIZE => '65536', REPLICATION_SCOPE => '0'}
{NAME => 'lookup_transactions_family', BLOOMFILTER => 'ROW', VERSIONS => '1', IN_MEMORY => 'false', KEEP
_DELETED CELLS => 'FALSE', DATA_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER', COMPRESSION => 'NONE', MIN_V
ERSIONS => '0', BLOCKCACHE => 'true', BLOCKSIZE => '65536', REPLICATION_SCOPE => '0'}
2 row(s) in 0.3090 seconds

hbase(main):002:0>
```

3. Altering "cc\_lookup\_data\_hive" table and setting VERSIONS to 10 for lookup\_transaction\_family since as per the problem statement we have to store last 10 transactions in lookup table.

```
alter 'cc_lookup_data_hive', {NAME => 'lookup_transactions_family', VERSIONS => 10}
```

```
hbase(main):001:0> describe 'cc_lookup_data_hive'
Table cc_lookup_data_hive is ENABLED
cc_lookup_data_hive
COLUMN FAMILIES DESCRIPTION
(NAME => 'lookup_card family', BLOOMFILTER => 'ROW', VERSIONS => '1', IN MEMORY => 'false', KEEP_DELETED
_CELLS => 'FALSE', DATA_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER', COMPRESSION => 'NONE', MIN_VERSIONS
=> '0', BLOCKCACHE => 'true', BLOCKSIZE => '65536', REPLICATION_SCOPE => '0')
(NAME => 'lookup_transactions family', BLOOMFILTER => 'ROW', VERSIONS => '1', IN MEMORY => 'false', KEEP
_DELETED_CELLS => 'FALSE', DATA_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER', COMPRESSION => 'NONE', MIN_V
ERSIONS => '0', BLOCKCACHE => 'true', BLOCKSIZE => '65536', REPLICATION_SCOPE => '0')
2 row(s) in 0.3090 seconds

hbase(main):002:0> alter 'cc_lookup_data_hive', {NAME => 'lookup_transactions_family', VERSIONS => 10}
Updating all regions with the new schema...
1/1 regions updated.
Done.
0 row(s) in 2.4090 seconds

hbase(main):003:0> █
```

4. Verify details of “cc\_lookup\_data\_hive” (hive-hbase integrated) table after altering version to 10 :  
**describe 'cc\_lookup\_data\_hive'**

```
[root@ip-172-31-21-198 ec2-user]# describe 'cc_lookup_data_hive'
bash: describe: command not found
[root@ip-172-31-21-198 ec2-user]# hbase shell
HBase Shell
Use "help" to get list of supported commands.
Use "exit" to quit this interactive shell.
Version 1.4.13, rUnknown, Fri Apr 17 15:18:24 UTC 2020

hbase(main):001:0> describe 'cc_lookup_data_hive'
Table cc_lookup_data_hive is ENABLED
cc_lookup_data_hive
COLUMN FAMILIES DESCRIPTION
(NAME => 'lookup_card family', BLOOMFILTER => 'ROW', VERSIONS => '1', IN MEMORY => 'false', KEEP_DELETED
_CELLS => 'FALSE', DATA_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER', COMPRESSION => 'NONE', MIN_VERSIONS
=> '0', BLOCKCACHE => 'true', BLOCKSIZE => '65536', REPLICATION_SCOPE => '0')
(NAME => 'lookup_transactions family', BLOOMFILTER => 'ROW', VERSIONS => '10', IN MEMORY => 'false', KE
P_DELETED_CELLS => 'FALSE', DATA_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER', COMPRESSION => 'NONE', MIN_
VERSIONS => '0', BLOCKCACHE => 'true', BLOCKSIZE => '65536', REPLICATION_SCOPE => '0')
2 row(s) in 0.2970 seconds

hbase(main):002:0> █
```