Creating Lookup Table

Purpose of document: This is a pdf document intent to contain creation of look-up table with columns specified in the problem statement in LogicMid.pdf.

Command to create the Lookup Table:

Create look_up_table in hbase:

echo "create 'look_up_table', 'card_details', 'Member_details', 'Location', 'Rule_params'" | hbase shell -n

```
[hadoop@ip-172-31-64-25 ~]$ echo "create 'look_up_table', 'card_details', 'Membe
r_details', 'Location', 'Rule_params'" | hbase shell -n
```

Create hive integrated look_uo_table:

CREATE EXTERNAL TABLE IF NOT EXISTS look_up_table (
  card_id STRING,
  card_purchase_dt STRING,
  transaction_dt STRING,
  member_id STRING,
  member_joining_dt STRING,
  country STRING,
  city STRING,
  UCL DOUBLE,
  postcode STRING,
  score INT
)
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ','
STORED BY "org.apache.hadoop.hive.hbase.HBaseStorageHandler"
WITH SERDEPROPERTIES ("hbase.columns.mapping" = "card_details:card_purchase_dt,
card_details:transaction_dt,Member_details:member_id, Member_details:member_joining_dt,
Location:country,Location:city,Rule_params:UCL,Rule_params:postcode, Rule_params:score")
TBLPROPERTIES ("hbase.table.name" = "look_up_table");

```
hive> CREATE EXTERNAL TABLE IF NOT EXISTS look_up_table (
    >     card_id STRING,
    >     card_purchase_dt STRING,
    >     transaction_dt STRING,
    >     member_id STRING,
    >     member_joining_dt STRING,
    >     country STRING,
    >     city STRING,
    >     UCL DOUBLE,
    >     postcode STRING,
    >     score INT
    > )
    > ROW FORMAT DELIMITED
    > FIELDS TERMINATED BY ','
    > STORED BY "org.apache.hadoop.hive.hbase.HBaseStorageHandler"
    > WITH SERDEPROPERTIES ("hbase.columns.mapping" = "card_details:card_purchase_dt, card_details:transaction_dt,Member_details:m
ember_id, Member_details:member_joining_dt, Location:country,Location:city,Rule_params:UCL,Rule_params:postcode, Rule_params:score
")
    > TBLPROPERTIES ("hbase.table.name" = "look_up_table");
OK
Time taken: 2.453 seconds
hive>
```

Note: Hive integrated look_up_table created in hbase

Command to see the table created:

describe 'look_up_table'

Screenshot of the created table

Screenshot of the table created:

```
hbase(main):001:0> describe 'look_up_table'
Table look_up_table is ENABLED
look_up_table
COLUMN FAMILIES DESCRIPTION
{NAME => 'Location', BLOOMFILTER => 'ROW', VERSIONS => '1', IN_MEMORY => 'false', KEEP_DELETED_CELLS => 'FALSE', DATA_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER', COM
PRESSION => 'NONE', MIN_VERSIONS => '0', BLOCKCACHE => 'true', BLOCKSIZE => '65536', REPLICATION_SCOPE => '0'}
{NAME => 'Member_details', 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 => 'Rule_params', 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 => 'card_details', 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'}
4 row(s) in 0.3170 seconds

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

Capstone project – Credit Card Fraud Detection (mid submission)

by

Susil Patro, Krishna Mohan & Ashmeet Singh Deol