

Loading Historical Transactions Data into NoSQL Database

Commands to load the past transactions data into NoSQL database

Code to load all CSV into NoSQL in batch mode.

```
import happybase

#create connection connection = happybase.Connection('localhost',
port=9090 ,autoconnect=False)

def open_connection():
connection.open()

def close_connection():
connection.close()

list all tables in Hbase

def list_tables():

print "fetching all table" open_connection() tables = connection.tables()
close_connection()

print "all tables fetched" return tables

#Creating the required table

def create_table(name,cf):

print "creating table " + name tables = list_tables() if name not in tables:
open_connection() connection.create_table(name, cf) close_connection()
print "table created" else: print "table already present"

#Get the pointer to a table def get_table(name):

open_connection() table = connection.table(name) close_connection()
return table
```

```
#batch insert data in events table def
batch_insert_data(filename,tableName):

print "starting batch insert of events" file = open(filename, "r") table =
get_table(tableName) open_connection() i=0 for line in file: temp =
line.strip().split(",")

#Skip the first row

if temp[0]!='card_id': table.put(bytes(i) , { 'info:card_id':bytes(temp[0]),
i=i+1 file.close()

print "batch insert done" 'info:member_id':bytes(temp[1]),
'info:amount':bytes(temp[2]), 'info:postcode':bytes(temp[3]),
'info:pos_id':bytes(temp[4]), 'info:transaction_dt':bytes(temp[5]),
'info:status':bytes(temp[6])}) close_connection() # Batch insert data of
card_transactions.csv file. create_table('card_transactions', {'info' :
dict(max_versions=5) })
batch_insert_data('card_transactions.csv','card_transactions')
```

<list the table in which the data is loaded to get the count of the rows of the table>

- 1) Login to Putty and enter as root user. 2) Run thrift server 3) Give Hbase Shell and press enter. 4) Give “list” to see all tables in hbase.

```
hbase(main):001:0> list
TABLE
card_transactions
employee
look_up_table
3 row(s) in 0.1970 seconds
=> ["card_transactions", "employee", "look_up_table"]
hbase(main):002:0> █
```

Give “ count ‘card_transactions’ “ in command line.

Screenshot of the table created 1) Scan ‘card_transactions’

```
=> ["card_transactions", "employee", "look_up_table"]
hbase(main):002:0> count 'card_transactions'
Current count: 1000, row: 10898
```

```
Current count: 53000, row: 9735
53292 row(s) in 4.1730 seconds
```

9736	column=info:pos_id, timestamp=1607957944969, value=546062979209182
9736	column=info:postcode, timestamp=1607957944969, value=95991
9736	column=info:status, timestamp=1607957944969, value=GENUINE
9736	column=info:transaction_dt, timestamp=1607957944969, value=03-01-2018 12:46:58
9737	column=info:amount, timestamp=1607957944972, value=816027
9737	column=info:card_id, timestamp=1607957944972, value=6011706374151856
9737	column=info:member_id, timestamp=1607957944972, value=199243620982420
9737	column=info:pos_id, timestamp=1607957944972, value=546062979209182
9737	column=info:postcode, timestamp=1607957944972, value=95991
9737	column=info:status, timestamp=1607957944972, value=GENUINE
9737	column=info:transaction_dt, timestamp=1607957944972, value=03-01-2018 23:18:59
9738	column=info:amount, timestamp=1607957944973, value=4042654
9738	column=info:card_id, timestamp=1607957944973, value=6011706374151856
9738	column=info:member_id, timestamp=1607957944973, value=199243620982420
9738	column=info:pos_id, timestamp=1607957944973, value=599304777798828
9738	column=info:postcode, timestamp=1607957944973, value=96727
9738	column=info:status, timestamp=1607957944973, value=GENUINE
9738	column=info:transaction_dt, timestamp=1607957944973, value=03-12-2017 08:53:52
9739	column=info:amount, timestamp=1607957944974, value=4034933
9739	column=info:card_id, timestamp=1607957944974, value=6011706374151856
9739	column=info:member_id, timestamp=1607957944974, value=199243620982420
9739	column=info:pos_id, timestamp=1607957944974, value=068029537978614
9739	column=info:postcode, timestamp=1607957944974, value=14511
9739	column=info:status, timestamp=1607957944974, value=GENUINE
9739	column=info:transaction_dt, timestamp=1607957944974, value=05-11-2017 10:15:33
974	column=info:amount, timestamp=1607957932834, value=5093161
974	column=info:card_id, timestamp=1607957932834, value=4380357644954006
974	column=info:member_id, timestamp=1607957932834, value=012871894714576
974	column=info:pos_id, timestamp=1607957932834, value=114439686882626
974	column=info:postcode, timestamp=1607957932834, value=10919
974	column=info:status, timestamp=1607957932834, value=GENUINE
974	column=info:transaction_dt, timestamp=1607957932834, value=05-10-2017 01:06:28
9740	column=info:amount, timestamp=1607957944975, value=3173523
9740	column=info:card_id, timestamp=1607957944975, value=6011706374151856
9740	column=info:member_id, timestamp=1607957944975, value=199243620982420
9740	column=info:pos_id, timestamp=1607957944975, value=771819356785069
9740	column=info:postcode, timestamp=1607957944975, value=72143
9740	column=info:status, timestamp=1607957944975, value=GENUINE
9740	column=info:transaction_dt, timestamp=1607957944975, value=05-12-2017 16:19:17
9741	column=info:amount, timestamp=1607957944976, value=2070998
9741	column=info:card_id, timestamp=1607957944976, value=6011706374151856
9741	column=info:member_id, timestamp=1607957944976, value=199243620982420
9741	column=info:pos_id, timestamp=1607957944976, value=771819356785069
9741	column=info:postcode, timestamp=1607957944976, value=72143