1. First we need to login to EMR and then switch to root user and run MySQL connector code

   sudo -i
   wget https://de-mysql-connector.s3.amazonaws.com/mysql-connector-java-8.0.25.tar.gz
   tar -xvf mysql-connector-java-8.0.25.tar.gz
   cd mysql-connector-java-8.0.25/
   sudo cp mysql-connector-java-8.0.25.jar /usr/lib/sqoop/lib/

2. Command to ingest the data from RDS into the HBase Table

sqoop import --connect jdbc:mysql://database-using.cb0sqm06ygfv.us-east-1.rds.amazonaws.com:3306/taxi \
--username admin --password admin123 \
--table yellow_taxi --hbase-table yellow_taxi_hbase --column-family cf1 --hbase-create-table \ --hbase-row-key tpep_pickup_datetime,tpep_dropoff_datetime --hbase-bulkload --split-by payment_type

- This is a sqoop command to import data form MySQL to Hbase table 'yellow_taxi_hbase'.
- The '—connect' is for connection for MySQL DB.
- Then '—username ' and '—password' are both used connect MySQL DB. '—table' to specify the table name in MySQL table.
- '—hbase-table' specifies the name of the table to import to.
- '—column-family' specific the name of column family
- '—hbase-create-table' create hbase table
- '—hbase-row-key' specifies the column that will be specified as row key in HBase

```
EEEEEEEEEEEEEEEEEEEEE MMMMMMM          MMMMMMM RRRRRRRRRRRRRRR
E::::::::::::::::::E M::::::M          M::::::M R::::::::::::::R
EE:::::EEEEEEEEE:::E M:::::::M        M:::::::M R:::::RRRRRR:::::R
  E::::E       EEEEE M::::::::M      M::::::::M RR::::R      R::::R
  E::::E             M:::::::::M    M:::::::::M   R:::R      R::::R
  E:::::EEEEEEEEEE    M:::::M:::M  M:::M:::::M   R:::RRRRRR:::::R
  E::::::::::::::E    M:::::M M:::M:::M M:::::M   R:::::::::::RR
  E:::::EEEEEEEEEE    M:::::M  M:::M:M  M:::::M   R:::RRRRRR::::R
  E::::E             M:::::M   M:::M   M:::::M   R:::R      R::::R
  E::::E       EEEEE M:::::M    MMM    M:::::M   R:::R      R::::R
EE:::::EEEEEEEE::::E M:::::M           M:::::M   R:::R      R::::R
E::::::::::::::::::E M:::::M           M:::::M RR::::R      R::::R
EEEEEEEEEEEEEEEEEEEEE MMMMMMM          MMMMMMM RRRRRRR      RRRRRR

[root@ip-172-31-65-22 ~]# hbase shell
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/lib/hadoop/lib/slf4j-reload4j-1.7.36.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/lib/hbase/lib/client-facing-thirdparty/slf4j-reload4j-1.7.33.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Reload4jLoggerFactory]
HBase Shell
Use "help" to get list of supported commands.
Use "exit" to quit this interactive shell.
For Reference, please visit: http://hbase.apache.org/2.0/book.html#shell
Version 2.4.15-amzn-0, rUnknown, Tue Feb 21 18:23:24 UTC 2023
Took 0.0025 seconds
hbase:001:0> list
TABLE
yellow_taxi_hbase
1 row(s)
Took 0.5947 seconds
=> ["yellow_taxi_hbase"]
hbase:002:0> describe yellow_taxi_hbase
Traceback (most recent call last):
NameError (undefined local variable or method `yellow_taxi_hbase' for #<HBaseReceiver:0x41582899>)
hbase:003:0> describe yellow_taxi_hbase;
Traceback (most recent call last):
NameError (undefined local variable or method `yellow_taxi_hbase' for #<HBaseReceiver:0x41582899>)
hbase:004:0> describe 'yellow_taxi_hbase'
Table yellow_taxi_hbase is ENABLED
yellow_taxi_hbase
COLUMN FAMILIES DESCRIPTION
{NAME => 'cf1', BLOOMFILTER => 'ROW', IN_MEMORY => 'false', VERSIONS => '1', KEEP_DELETED_CELLS => 'FALSE', DATA_BLOCK_ENCODING => 'NONE', COMPRESSION => 'NONE', TTL => 'FOREVER', MIN_VERSI
ONS => '0', BLOCKCACHE => 'true', BLOCKSIZE => '65536', REPLICATION_SCOPE => '0'}

1 row(s)
Quota is disabled
Took 0.1934 seconds
hbase:005:0>
```