

Data imported into NoSql table using sqoop job-
amazon_hbase_card_transactions_new

Data imported into hive tables using sqoop job-
card_member_hive_new and member_score_hive_1

Since for the batch job we needed to upload the data of card_transactions only once, so I did not include it in my coordinator job to upload it every 4 hours. For the remaining two tables from RDS, I created sqoop jobs in coordinator to upload data from them every 4 hours.

Once I have the data in HDFS, I ran HIVE scripts on the tables to create multiple temp tables to calculate, latest txn date, post code, standard deviation, moving average, ucl seperately.

After generating these temp tables which have corresponding values for each unique card id, I used them to create the final look-up table by left joining them.

The final lookup table has all the required data.

****Please note - I stored all the queries mentioned in "PreAnalysis.txt" file in my SQL file called queries.sql**