**Credit\_Card\_System Case Study By Sureshkumar Packiyarajah**

**Batch:DE01**

**Date: 05/03/2018**

1. **Core Java:**

File name: CreditCard\_System.rar

JDBC Connection username: root

Password: Oracle

IP/Localhost :127.0.0.1

Software/Platform Used: eclipse oxygen under Ubuntu.

2. **RDBMS/MySQL Description**:

File name: db.rar

It has three files: 1.CDW\_SAPP\_CDW\_SAPP\_BRANCH.sql

2.CDW\_SAPP\_CDW\_SAPP\_CREDITCARD.sql

3.CDW\_SAPP\_CDW\_SAPP\_CUSTOMER.sql

3. **Hadoop/HDFS/Dataware housing:**

File name: Hadoop.txt

Hortonworks username: root

Password: hadoop

IP/Localhost :127.0.0.1

Software/Platform Used: Hortonworks HDP 2.5 in Virtulbox.

Connection String used: jdbc:mysql://127.0.0.1:3306/CDW\_SAPP --username root --password hadoop --driver com.mysql.jdbc.Driver

It has three Four Steps: 1.Customer\_Sqoop

2. Branch\_Sqoop

3. CreditCard\_Sqoop

4. Timeid\_Sqoop

4. **Hive and Partition:**

File name: Hive.txt

I have done Customer PARTITION in Hive table and I used Branch, CreditCard and Timeid remaining same as Internal and External Table.

Location: /user/maria\_dev/Credit\_Card\_System/

5. **Oozie (Sqoop and Hive):**

File name: sqoop\_etl.txt

Hortonworks username: root

Password: hadoop

IP/Localhost :127.0.0.1

Software/Platform Used: Hortonworks HDP 2.5 in Virtulbox.

6. **Oozie (Sqoop and Hive Optimized):**

File name: Oozie\_Jobs.rar

It has jobs.properties, coordinator.xml, workflow.xml also internal.hive and external.hive files.

7. **Hive Visualization:**

File name: hive\_vizualization.rar

It has snapshot of visualization what I did thought two queries.

Q1) Find the top 20 zip codes(hint: branch\_zip) by total transaction value for each transaction type?

Q2) Find total transaction value for each transaction type by Quarter in 2018?