**Core Java**

* Project name is CreditCardSystem.
* Open project in your IDE and be sure to change the database parameters.
  + In database package go to DatabaseConnection.java file and change database name as well as username and password if needed.
* Clean and Build project
* Run myquery.java file

**RDBMS/mySQL** **Description** :

* To create tables and insert data for this project, run following mySQL file:
  + CDW\_SAPP\_CDW\_SAPP\_BRANCH
  + CDW\_SAPP\_CDW\_SAPP\_CREDIITCARD
  + CDW\_SAPP\_CDW\_SAPP\_CUSTOMER
* This file will create database for you named CDW\_SAPP

**Hadoop/hdfs/dataware housing**

* In HDFS we are using **maria\_dev** user. Therefore data transferred from relational database with Sqoop is stored in /user/maria\_dev/Credit\_Card\_System

**Hive and Partition**

* Hive imports consist of four tables (Branch, Credit Card, Time and Customer). Each table is external table that loads data files imported by sqoop.
* **CreateBranchTable.sql**
  + This file creates Branch tables called CDW\_SAPP\_D\_BRANCH
  + Loads data from /user/maria\_dev/Credit\_Card\_System/Branch
* **CreateCreditCardTable.sql**
  + This file creates CreditCrad tables called CDW\_SAPP\_F\_CREDITCARD
  + Load data from/user/maria\_dev/Credit\_Card\_System/CreditCard
* **CreateCustomerTable.sql**
  + This file creates Customer tables called CDW\_SAPP\_D\_CUSTOMER
  + Loads data from /user/maria\_dev/Credit\_Card\_System/Customer
* **CreateTimeTable.sql**
  + This file creates Time tables called CDW\_SAPP\_D\_TIME
  + Loads data from /user/maria\_dev/Credit\_Card\_System/TimeID
* **Incrementals.sql**
  + This file insert new data in incremental way.
  + It is used in oozie with coordinators for incremental update.

**Oozie (Sqoop and Hive)**

* Before running any sqoop jobs it is important to run sqoop metastore.
* In /SqoopImport directory there is file called **sqoopjobs.sh** which is shell script that will create all Sqoop jobs.
* Transfer **sqoopjobs.sh** file to your local path
* Run following command as root user: root

/user/maria\_dev/sqoopjobs.sh

* Run it: ./sqoopjobs.sh
* **Transfer directory /OozieWorkflow and /HiveImports to both local and hdfs path:**
  + - Type: hadoop fs --put OozieWorkflow/ /user/maria\_dev/
    - Type: hadoop fs --put HiveImports/ /user/maria\_dev/
* --Upload java-json.jar file:
  + **oo** There will be file java-json.jar
  + **oo** Use Ambari to upload that file to /user/oozie/share/lib/ **lib\_\*\*\*\*\*\*\*** /sqoop/
  + **oo** Change lib\_\*\*\*\*\*\*\* to your directory name
* --Before running Oozie import mysql database required for this project. File to insert is in /mySQL/CDW\_SAPP.sql
* Run this command to start Oozie to create Hive tables and import data from relational database:
  + **oo** oozie job --oozie <http://localhost:11000/oozie> -config /home/maria\_dev/OozieWorkflow/Initialize/job.properties -run

**Oozie (Sqoop and Hive optimized)**

* After Initialized Oozie workflow, we can run Incremental Oozie workflow for incremental update
* We are running same sqoop jobs and one additional hive query for updating data with dynamic partitioning.
* To run optimized oozie run:
  + **oo** oozie job --oozie <http://localhost:11000/oozie> -config /home/maria\_dev/OozieWorkflow/Incremental/job.properties -run

**Visualization**

* In /HiveVisualization directory there are two hive querys.