**Core Java**

* Java project location is:

*2.1.1 Customer and Transaction Details Module\Core Java\Credit Card Management System.zip*

**RDBMS/mySQL Description**:

* MySQL database script location is:

*2.1.1 Customer and Transaction Details Module\RDBMS - MySQL\CDW\_SAPP.zip*

* You need to extract the zip file and execute the sql script from MySQL Workbench in order to recreate the database and all of its tables

**Hadoop/hdfs/dataware housing**

* *hadoop.txt* – this file cointains all the necessary commands that need to be executed to perform ET process
* File location is: *2.2.1 Data Extraction and Transportation with Sqoop\hadoop.txt*

**Hive and Partition**

* To perform data loading using Hive execute scripts from either *Ambari Hive* view or by executing the following command in terminal:

*hive –f “hdfs://sandbox.hortonworks.com:8020/user/maria\_dev/scripts/hive/scriptName.txt”* where scriptName.txt is the name of the script

* *branchTable.txt* - used for loading data into branch table
* *ccTable.txt -* used for loading data into credit card table
* *customerTable.txt –* used for loading data into customer table
* *timeTable.txt -* used for loading data into time table
* All the above file are located under: *2.2.2 Data Loading Module\ directory*
* Files need to be placed under: /*user/maria\_dev/scripts/hive/ directory on HDFS*

**Oozie (Sqoop and Hive)**

* *coordinator.xml* – coordinator file
* *dropTables.txt* – hive script for dropping tables
* *hadoop.txt* – in order to run the coordinator job, commands that are listed here need to be executed from the terminal
* *job.properties* – properties file that needs to be places under /home/maria\_dev/oozie/ directory on local filesystem
* *workflow.xml* – workflow file
* Zip file that contains all the above files is located here:

*2.2.3 Automating the Process with Oozie\oozie.zip*

* All the files except *jobs.properties* and *hadoop.txt* need to be placed under /*user/maria\_dev/scripts/oozie/* directory on HDFS
* *java-json.jar* file needs to be placed in this location: /user/maria\_dev/lib/ on HDFS

**Oozie (Sqoop and Hive optimized)**

* *coordinator.xml*
* *drop\_staging\_tables.txt*
* *hadoop.txt* – in order to run the coordinator job, commands that are listed here need to be executed from the terminal
* *insert\_incremental\_into\_branch.txt*
* *insert\_incremental\_into\_creditcard.txt*
* *insert\_incremental\_into\_customer.txt*
* *insert\_incremental\_into\_time.txt*
* *job.properties -* properties file that needs to be places under: /home/maria\_dev/oozie\_optimized/ directory on local filesystem
* *populate\_staging\_branch.txt*
* *populate\_staging\_customer.txt*
* *workflow.xml*
* Zip file that contains all the above files is located here:

*2.2.2 Data Loading Module\oozie\_optimized.zip*

* All the files except *jobs.properties* and *hadoop.txt* need to be placed under: /*user/maria\_dev/scripts/oozie\_optimized/ directory on HDFS*
* *java-json.jar* file needs to be placed in this location: /user/maria\_dev/lib/ on HDFS

**Visualization**

* File with hive queries and graphs is located here:

*2.2.5 Data Visualization\Data Visualization.docx*