**Requirements**

* These files were given to the class by the Product Owner in this case: Platform by Per Scholas.
* requirements directory
  + casetudy Flow (1.2).pdf
  + Credit Card Management System\_SRD.pdf
  + Functional Requirements.pdf: This file lists the requirements for the application.
  + Mapping Document.xlsx
  + Source File Structure.xlsx

**Core Java**

* These files contain the java code in an Eclipse project called CDW\_SAPP that can be imported into Eclipse and run.
* 2.1.1\_2.1.2\_java\_eclipse\_project
  + CDW\_SAPP
    - src\com\cdw\dao: Data Access Objects
      * AbstractDAO.java
      * CustomerDAO.java
      * TransactionDAO.java
    - src\com\cdw\model
      * Customer.java
      * Transaction.java
    - src\com\cdw\resources
      * db.properties
      * Queries.java
    - src\com\cdw\runner
      * CustomerMenu.java
      * MainMenu.java: contains main method
      * Menu.java: Abstract class
      * TransactionMenu.java
      * TransactionTypeMenu.java

**RDBMS/mySQL**

* These files contain the schema and data of the MySQL database.
* 2.1.1\_2.1.2\_mysql
  + CDW\_SAPP.sql: schema for database
  + CDW\_SAPP\_BRANCH.csv: data to be imported into corresponding table
  + CDW\_SAPP\_CREDITCARD.csv: data to be imported into corresponding table
  + CDW\_SAPP\_CUSTOMER.csv: data to be imported into corresponding table

**Hadoop/hdfs/dataware housing**

* This file contains commands that are manually run to import data from mysql to hdfs and then load into hive data warehouse.
* 2.2.1\_2.2.2\_sqoop\_hive
  + 2.2.1\_2.2.2.txt: mysql, hadoop, sqoop, and hive commands

**Oozie (Sqoop and Hive)**

* These files use oozie to automate the running of the sqoop and hive commands. These files delete all data imported into hdfs and imports the complete data set from mysql with each run.
* 2.2.3\_oozie
  + case\_study\_2.2.3\_notes.txt: hive, hadoop, and oozie commands that are manually run
  + coordinator.xml: file needs to be put in hdfs /user/maria\_dev/cdw\_sapp/oozie/ directory; run by job\_coordinator.properties
  + external.hive: hive file that creates external tables
  + job.properties: file needs to be put in vm: /root/Documents/case\_study/oozie/
  + job\_coordinator.properties: file needs to be put in vm: /root/Documents/case\_study/oozie/
  + loaddata.hive: hive file that loads data into external tables
  + workflow.xml: file needs to be put in hdfs /user/maria\_dev/cdw\_sapp/oozie/ directory; run by job.properties or coordinator.xml; contains sqoop and hive actions

**Oozie (Sqoop and Hive optimized)**

* These files use oozie to automate the running of the sqoop and hive commands. These files only import new data into hdfs from mysql with each run.
* 2.2.3\_oozie\_optimized
  + case\_study\_2.2.4\_notes.txt: bash, hadoop, and oozie commands that are manually run
  + coordinator.xml: file needs to be put in hdfs /Credit\_Card\_System\_224/oozie/ directory; run by job\_coordinator.properties
  + external.hive: hive file that creates external tables; these commands are manually run once before oozie coordinator job is started
  + job.properties: file needs to be put in vm: /root/Documents/case\_study/oozie224/
  + job\_coordinator.properties: file needs to be put in vm: /root/Documents/case\_study/oozie224/
  + loaddata.hive: hive file that loads data into external tables
  + sqoop\_create\_jobs.sh: these commands are manually run once before oozie coordinator job is started
  + workflow.xml: file needs to be put in hdfs /Credit\_Card\_System\_224/oozie/ directory; run by job.properties or coordinator.xml; contains sqoop actions

**Visualization**

* These files are run in Ambari Hive View.
* 2.2.5\_hive visualization
  + 2.2.5.txt: The queries and the result set data
  + 2.2.5\_1.png: A graph created using hive visualization for the first query
  + 2.2.5\_2.png: A graph created using hive visualization for the second query