This the READ ME document for CASE STUDY -2.2 Functional Requirements – ETL of Data

The package is consist of 5 files.

FIRST PART of FUNCTION REQUIREMENT-

**Data Extraction and Transportation with Sqoop**

1. Sqoop Queries 2.2.1.DOC-

This contains the queries of SQOOP to run the SQOOP jobs which pulls the data from MYSQL to HDFS.

SECOND PART of FUNCTION REQUIREMENT-

**Data Loading with Hive**

1. *HIVE QUERIES for table CDW\_SAPP\_CUSTOMER.DOC*

This contains the HIVE queries of CUSTOMER TABLE which is creating Temp Table , Transformation Table and Partition Table .

1. *HIVE QUERIES for table CDW\_SAPP\_BRANCH.DOC*

This contains the HIVE queries of BRANCH TABLE which is creating Temp Table , Transformation Table and Partition Table .

1. *HIVE QUERIES for table CDW\_SAPP\_CREDITCARD.DOC*

This contains the HIVE queries of CREDITCARD TABLE which is creating Temp Table , Transformation Table and Partition Table .

1. *HIVE QUERIES for table CDW\_SAPP\_TIMEID.DOC*

This contains the HIVE queries of TIMEID TABLE which is creating Temp Table , Transformation Table and Partition Table .

All the SQOOP queries are ran using PUTTY Terminals and All the HIVE queries are ran using AMBARI HIVE view.