**Extraction: Importing tables from Oracle to Hive**

**Successful Sqoop Import:** Grade, Trainer, Address, Category,

Note, Batch, Trainee, Panel, Panel feedback, Assessment

**Tables to be made:** Assignment, Task, Task completion

**Requirements before running sqoop commands**

1. **Download OJDBC6 jar from oracle**
2. **Copy into sqoop: sudo mv ojdbc6.jar /var/lib/sqoop**
3. **May have to create a spark job that will cleanse data for plicrow and other symbols**
   1. **May have to create spark for one or all files**
      1. **Octal symbols to check invalid symbols**
      2. **Will get back to you on that**
4. **Please create CALIBER database in hue before running sqoop commands**
5. **SQOOP 2 is deprecated, so you don’t have to turn it on. The jar will do the work**
6. **NOTE: Some symbols may mess up the column integrity when you sqoop import**
   1. **I.E.) CALIBER\_NOTE HAS THIS ISSUE WITH THE PARAGRAPH SYMBOL AT NOTE\_ID = 6308**

**(Testing connection with data and testing with query)**

sqoop eval --connect jdbc:oracle:thin:@caliber-test.chueiwozbnfz.us-east-1.rds.amazonaws.com:1521/orcl --username associate --password rev123 --query "SELECT \* FROM CALIBER.CALIBER\_TRAINEE"

**(Category table)**

sqoop import --connect jdbc:oracle:thin:@caliber-test.chueiwozbnfz.us-east-1.rds.amazonaws.com:1521/orcl --username associate --password rev123 --table CALIBER.CALIBER\_CATEGORY --hive-import --create-hive-table --hive-table CALIBER.CALIBER\_CATEGORY

**(For Trainer table)**

sqoop import --connect jdbc:oracle:thin:@caliber-test.chueiwozbnfz.us-east-1.rds.amazonaws.com:1521/orcl --username associate --password rev123 --table CALIBER.CALIBER\_TRAINER --hive-import --create-hive-table --hive-table CALIBER.CALIBER\_TRAINER

**V2- MOST UPDATED**

sqoop import --connect jdbc:oracle:thin:@caliber-test.chueiwozbnfz.us-east-1.rds.amazonaws.com:1521/orcl --username associate --password rev123 --table CALIBER.CALIBER\_TRAINER --fields-terminated-by \~ --target-dir /user/cloudera/CALIBER/CALIBER\_TRAINER -m 1 --hive-drop-import-delims

**(For Trainee table)**

sqoop import --connect jdbc:oracle:thin:@caliber-test.chueiwozbnfz.us-east-1.rds.amazonaws.com:1521/orcl --username associate --password rev123 --table CALIBER.CALIBER\_TRAINEE --hive-import --create-hive-table --hive-table CALIBER.CALIBER\_TRAINEE

**(Address table)**

sqoop import --connect jdbc:oracle:thin:@caliber-test.chueiwozbnfz.us-east-1.rds.amazonaws.com:1521/orcl --username associate --password rev123 --table CALIBER.CALIBER\_ADDRESS --hive-import --create-hive-table --hive-table CALIBER.CALIBER\_ADDRESS

**(Grade table)**

sqoop import --connect jdbc:oracle:thin:@caliber-test.chueiwozbnfz.us-east-1.rds.amazonaws.com:1521/orcl --username associate --password rev123 --table CALIBER.CALIBER\_GRADE --hive-import --create-hive-table --hive-table CALIBER.CALIBER\_GRADE

**V2- MOST UPDATED**

sqoop import --connect jdbc:oracle:thin:@caliber-test.chueiwozbnfz.us-east-1.rds.amazonaws.com:1521/orcl --username associate --password rev123 --table CALIBER.CALIBER\_GRADE --fields-terminated-by \~ --target-dir /user/cloudera/CALIBER/CALIBER\_GRADE -m 1 --hive-drop-import-delims --hive-overwrite

**(Note table) - jilrm, HD, csl, lf, jl**

**REGEX = STRING COLUMNS, TABLE FROM HDFS**

Sqoop into HDFS

sqoop import --connect jdbc:oracle:thin:@caliber-test.chueiwozbnfz.us-east-1.rds.amazonaws.com:1521/orcl --username associate --password rev123 --table CALIBER.CALIBER\_NOTE --fields-terminated-by “,” --target-dir /user/cloudera/CALIBER -m 1

V2 - **CURRENT USE MOST EFFECTIVE**

sqoop import --connect jdbc:oracle:thin:@caliber-test.chueiwozbnfz.us-east-1.rds.amazonaws.com:1521/orcl --username associate --password rev123 --table CALIBER.CALIBER\_NOTE --fields-terminated-by \~ --target-dir /user/cloudera/CALIBER/CALIBER\_NOTE -m 1 --hive-drop-import-delims

NO FIELDS TERM

sqoop import --connect jdbc:oracle:thin:@caliber-test.chueiwozbnfz.us-east-1.rds.amazonaws.com:1521/orcl --username associate --password rev123 --table CALIBER.CALIBER\_NOTE --hive-import --create-hive-table --hive-table CALIBER.CALIBER\_NOTE --hive-overwrite -m 1

FIELDS TERM

sqoop import --connect jdbc:oracle:thin:@caliber-test.chueiwozbnfz.us-east-1.rds.amazonaws.com:1521/orcl --username associate --password rev123 --table CALIBER.CALIBER\_NOTE --fields-terminated-by , --hive-import --create-hive-table --hive-table CALIBER.CALIBER\_NOTE --hive-overwrite -m 1

FIELDS TERM - V2

sqoop import --connect jdbc:oracle:thin:@caliber-test.chueiwozbnfz.us-east-1.rds.amazonaws.com:1521/orcl --username associate --password rev123 --table CALIBER.CALIBER\_NOTE --fields-terminated-by “,” --hive-import --hive-overwrite --create-hive-table --hive-table CALIBER.CALIBER\_NOTE --hive-drop-import-delims

NON HIVE IMPORT

sqoop import --connect jdbc:oracle:thin:@caliber-test.chueiwozbnfz.us-east-1.rds.amazonaws.com:1521/orcl --username associate --password rev123 --table CALIBER.CALIBER\_NOTE --fields-terminated-by , --hive-import --create-hive-table --hive-table CALIBER.CALIBER\_NOTE --hive-overwrite -m 1

SQOOP IMPORT DIRECT ORACLE TO HDFS

sqoop import --direct --connect jdbc:oracle:thin:@caliber-test.chueiwozbnfz.us-east-1.rds.amazonaws.com:1521/orcl --username associate --password rev123 --table CALIBER.CALIBER\_NOTE -m 1

**(Batch table)**

sqoop import --connect jdbc:oracle:thin:@caliber-test.chueiwozbnfz.us-east-1.rds.amazonaws.com:1521/orcl --username associate --password rev123 --table CALIBER.CALIBER\_BATCH --hive-import --create-hive-table --hive-table CALIBER.CALIBER\_BATCH

**(Panel table)**

sqoop import --connect jdbc:oracle:thin:@caliber-test.chueiwozbnfz.us-east-1.rds.amazonaws.com:1521/orcl --username associate --password rev123 --table CALIBER.CALIBER\_PANEL --hive-import --create-hive-table --hive-table CALIBER.CALIBER\_PANEL

**(Panel Feedback table)**

sqoop import --connect jdbc:oracle:thin:@caliber-test.chueiwozbnfz.us-east-1.rds.amazonaws.com:1521/orcl --username associate --password rev123 --table CALIBER.CALIBER\_PANEL\_FEEDBACK --hive-import --create-hive-table --hive-table CALIBER.CALIBER\_PANEL\_FEEDBACK

**(Assessment table)**

sqoop import --connect jdbc:oracle:thin:@caliber-test.chueiwozbnfz.us-east-1.rds.amazonaws.com:1521/orcl --username associate --password rev123 --table CALIBER.CALIBER\_ASSESSMENT --hive-import --create-hive-table --hive-table CALIBER.CALIBER\_ASSESSMENT

**(Task table - DOESN’T EXIST - MUST MAKE TABLE - NO SQOOP JOB)**

sqoop import --connect jdbc:oracle:thin:@caliber-test.chueiwozbnfz.us-east-1.rds.amazonaws.com:1521/orcl --username associate --password rev123 --table CALIBER.CALIBER\_TASK --hive-import --create-hive-table --hive-table CALIBER.CALIBER\_TASK

**(Task Completion - DOESN’T EXIST - MUST MAKE TABLE - NO SQOOP JOB)**

sqoop import --connect jdbc:oracle:thin:@caliber-test.chueiwozbnfz.us-east-1.rds.amazonaws.com:1521/orcl --username associate --password rev123 --table CALIBER.CALIBER\_TASK\_COMPLETION --hive-import --create-hive-table --hive-table CALIBER.CALIBER\_TASK\_COMPLETION