#Author: Abuchi Okeke

#Date: 16/10/2020

# Code process for Hive and HBase Assignment

**#1 Create HDFS Directory**

hdfs dfs -mkdir /user/databases/mysql

hdfs dfs -mkdir /user/databases/sqlserver

hdfs dfs -mkdir /user/databases/postgresql

hdfs dfs -mkdir /user/databases/csv

**#2 Execute Sqoop**

./sqoop.sh

**#3 Move a .csv file to gcp using scp command or sftp**

scp data.csv Buchi@35.225.142.191:~

**#4 Establish sftp connection with gcp vm instance**

sftp Buchi@35.225.142.191

sftp> put data.csv

**#5 Get file from gcp to LFS**

sftp> get spotify\_datasets.csv

**#Upload to hdfs**

hdfs dfs -put ~/datasets/spotify\_datasets.csv /user/databases/csv/opensource

**#Execute scripts**

hive -f create\_hive\_internal\_tables.hql

hive -f create\_hive\_external\_tables.hql

hive -f create\_hive\_hbase\_tables.hql

**Screenshots**

1. Raw\_Data\_On\_HDFS

Graphical user interface, table

Description automatically generated

1. Tables on hdfs

Graphical user interface

Description automatically generated

1. SFTP

Text

Description automatically generated

1. Raw, dsl and asl on production server

Text

Description automatically generated

1. Raw (external tables)

Text

Description automatically generated

1. DSL (Internal tables)

A screenshot of a computer

Description automatically generated

1. ASL (application layer)

A screenshot of a computer

Description automatically generated

1. Hive – Hbase Integration

A screenshot of a computer

Description automatically generated