Sqoop assignments (TLO 1.1)

1. Grant privileges to dbs in mysql for users to connect from SQOOP Implement the following in SQOOP

grant all privileges on *.* to 'msis'@'localhost' identified by 'msis@123';

2. List all databases in mysql through SQOOP

sqoop list-databases --connect jdbc:mysql://localhost?useSSL=false --username msis --password msis@123

3. List all tables in selected database

sqoop list-tables --connect jdbc:mysql://localhost/Programs?useSSL=false --username msis --password msis@123

4. Import table into HDFS using target-dir method

sqoop import --connect jdbc:mysql://localhost/retail_db?useSSL=false --username msis --password msis@123 --table customers --target-dir '/testcustomer'

sqoop import --connect jdbc:mysql://localhost/retail_db?useSSL=false --username msis --password msis@123 --table orders --target-dir '/testorders'

5. Import table into HDFS using warehouse-dir method

sqoop import --connect jdbc:mysql://localhost/retail_db?useSSL=false --username msis --password msis@123 --table orders --warehouse-dir '/retail_db'

6. To import partial data from specified table

sqoop import --connect jdbc:mysql://localhost/retail_db?useSSL=false --username msis --password msis@123 --table categories --target-dir '/testcategories' --where "category_id <30"

7. To import only data from selected columns of table

sqoop import --connect jdbc:mysql://localhost/retail_db?useSSL=false --username msis --password msis@123 --columns product_id,product_name,product_price --table products --target-dir '/testproductscolumns'

8. To import all tables from given database

sqoop import-all-tables --connect jdbc:mysql://localhost/retail_db?useSSL=false --username msis --password msis@123 --warehouse-dir '/testdatabase'

9. To exclude some tables while importing all tables

sqoop import-all-tables --connect jdbc:mysql://localhost/retail_db?useSSL=false --username msis --password msis@123 --exclude-tables table1,table2 --warehouse-dir '/testdatabase'

10. To export table from HDFS to local dB

sqoop export --connect jdbc:mysql://localhost/test_export?useSSL=false --username msis --password msis@123 --table orders --export-dir /retail_db/orders

Implement following using HIVE tool (TLO 1.1)

1. Find unique states in customer table

select distinct customer_state from customers;

2. Find number of customers from each state

select state, count (*) as num_customers from customers GROUP BY state;

3. List and count number of unique fnames from customer table

select count(DISTINCT fname) as num_unique_fnames , fname from customer GROUP BY fname;

4. List and count unique cities from customer table

select count (DISTINCT city) as num_unique_cities, city from customer GROUP BY city;

5. Find customer id who placed maximum orders

select customer_id FROM orders GROUP BY customer_id ORDER BY COUNT(*) DESC LIMIT 1;