Session 5

Task 1

Use Sqoop tool to export data present in SQOOPOUT folder made while demo of Import table

Ans:create table employee (firstname varchar(10), lastname varchar(10));

insert into employee values ("syed","rizvi");

Output in mysql shell:

```
mysql> select * from employee;
+-----------+----------+
| firstname | lastname |
+-----------+----------+
| syed      | rizvi    |
| syed      | rizvi    |
+-----------+----------+
2 rows in set (0.00 sec)
```

In Hdfs we use command :

hadoop fs -cat / sqoopout/part-m-0000 to get output (here we create directory in hdfs Sqoopout where putput is stored).

Output:

```
cat: /: Is a directory
[acadgild@localhost ~]$ hadoop fs-ls /
Error: Could not find or load main class fs-ls
[acadgild@localhost ~]$ hadoop fs -ls /
18/07/28 12:58:51 WARN util.NativeCodeLoader: Unable to load native-hadoop librar
iltin-java classes where applicable
Found 16 items
drwxr-xr-x   - acadgild supergroup          0 2018-07-12 14:11 /SqoopSimpliDbAll
drwxr-xr-x   - acadgild supergroup          0 2018-07-22 19:25 /delme
drwxr-xr-x   - acadgild supergroup          0 2018-07-07 19:10 /hadoop
drwxr-xr-x   - acadgild supergroup          0 2018-07-15 17:20 /hadoopdata
drwxr-xr-x   - acadgild supergroup          0 2018-07-12 17:14 /myoutput
drwxr-xr-x   - acadgild supergroup          0 2018-07-12 17:32 /myoutput_mean
drwxr-xr-x   - acadgild supergroup          0 2018-07-12 17:42 /myoutput_std
drwxr-xr-x   - acadgild supergroup          0 2018-07-12 17:38 /myoutputmean
drwxr-xr-x   - acadgild supergroup          0 2018-07-12 17:22 /myoutputmedian
drwxr-xr-x   - acadgild supergroup          0 2018-07-12 15:08 /output
drwxr-xr-x   - acadgild supergroup          0 2018-07-07 18:20 /sqoopout
drwxr-xr-x   - acadgild supergroup          0 2018-07-07 17:58 /test.txt
drwx------   - acadgild supergroup          0 2018-07-14 17:07 /tmp
drwxr-xr-x   - acadgild supergroup          0 2018-07-14 17:13 /user
drwxr-xr-x   - acadgild supergroup          0 2018-07-07 23:00 /useracadgild
drwxr-xr-x   - acadgild supergroup          0 2018-07-07 22:11 /word-count.txt
[acadgild@localhost ~]$ hadoop fs -ls /sqoopout
```

```
[acadgild@localhost ~]$ hadoop fs -ls /sqoopout
18/07/28 13:19:19 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using bu
iltin-java classes where applicable
Found 2 items
-rw-r--r--   1 acadgild supergroup          0 2018-07-07 18:20 /sqoopout/_SUCCESS
-rw-r--r--   1 acadgild supergroup         11 2018-07-07 18:20 /sqoopout/part-m-00000
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ hadoop fs -cat /sqoopout/part-m-00000
18/07/28 13:19:49 WARN util.NativeCodeLoader: Unable to load native-hadoop library                . using bu
iltin-java classes where applicable
syed,rizvi
[acadgild@localhost ~]$
```

Task2

Use Sqoop tool to export data present in SQOOPOUT folder made while demo of Import table

with parameter person_id =3

Ans: Create a table Person in mysql   shell and put some data in it .

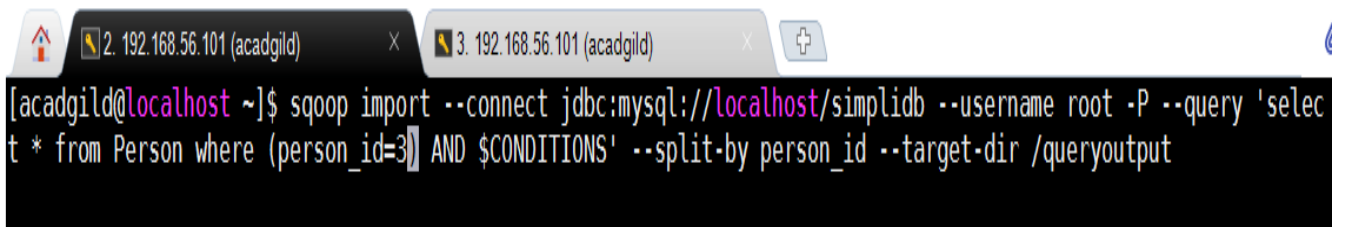To check data in person table use command:

Select * from Person;

output:

```
mysql> select * from Person;
+-----------+-------+-------+------------+-----------+
| person_id | lname | fname | area       | city      |
+-----------+-------+-------+------------+-----------+
|         1 | Rizvi | Syed  | Whitefield | Bangalore |
|         2 | Rizvi | Syed  | Whitefield | Bangalore |
|         3 | Rizvi | Syed  | Whitefield | Bangalore |
|         4 | Rizvi | Syed  | Whitefield | Bangalore |
|         5 | Rizvi | Syed  | Whitefield | Bangalore |
|         6 | Rizvi | Syed  | Whitefield | Bangalore |
|         7 | Rizvi | Syed  | Whitefield | Bangalore |
|         8 | Rizvi | Syed  | Whitefield | Bangalore |
|         9 | Rizvi | Syed  | Whitefield | Bangalore |
|        10 | Rizvi | Syed  | Whitefield | Bangalore |
|        11 | Rizvi | Syed  | Whitefield | Bangalore |
|        12 | Rizvi | Syed  | Whitefield | Bangalore |
|       789 | Rizvi | Syed  | Whitefield | Bangalore |
+-----------+-------+-------+------------+-----------+
13 rows in set (0.00 sec)
```

Command we use to get the output for person_id = 3

```
2. 192.168.56.101 (acadgild)      X    3. 192.168.56.101 (acadgild)      X

[acadgild@localhost ~]$ sqoop import --connect jdbc:mysql://localhost/simplidb --username root -P --query 'selec
t * from Person where (person_id=3) AND $CONDITIONS' --split-by person_id --target-dir /queryoutput
```

```
18/07/30 01:15:14 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
Mon Jul 30 01:15:21 IST 2018 WARN: Establishing SSL connection without server's identity verification is not re
ommended. According to MySQL 5.5.45+, 5.6.26+ and 5.7.6+ requirements SSL connection must be established by def
ult if explicit option isn't set. For compliance with existing applications not using SSL the verifyServerCerti
icate property is set to 'false'. You need either to explicitly disable SSL by setting useSSL=false, or set use
SL=true and provide truststore for server certificate verification.
18/07/30 01:15:22 INFO db.DBInputFormat: Using read commited transaction isolation
18/07/30 01:15:22 INFO db.DataDrivenDBInputFormat: BoundingValsQuery: SELECT MIN(person_id), MAX(person_id) FRO
 (select * from Person where (person_id=3) AND  (1 = 1) ) AS t1
18/07/30 01:15:22 INFO mapreduce.JobSubmitter: number of splits:1
18/07/30 01:15:23 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1532888272570_0007
18/07/30 01:15:26 INFO impl.YarnClientImpl: Submitted application application_1532888272570_0007
18/07/30 01:15:26 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1532888
72570_0007/
18/07/30 01:15:26 INFO mapreduce.Job: Running job: job 1532888272570 0007
```

```
        Map-Reduce Framework
                Map input records=1
                Map output records=1
                Input split bytes=109
                Spilled Records=0
                Failed Shuffles=0
                Merged Map outputs=0
                GC time elapsed (ms)=326
                CPU time spent (ms)=8190
                Physical memory (bytes) snapshot=162189312
                Virtual memory (bytes) snapshot=2086748160
                Total committed heap usage (bytes)=95420416
        File Input Format Counters
                Bytes Read=0
        File Output Format Counters
                Bytes Written=34
18/07/30 01:16:31 INFO mapreduce.ImportJobBase: Transferred 34 bytes in 77.38 seconds (0.4394 bytes/sec)
18/07/30 01:16:31 INFO mapreduce.ImportJobBase: Retrieved 1 records.
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$
```

For the output check we use command:

  Hadoop fs -ls /queryoutput

```
[acadgild@localhost ~]$ hadoop fs -ls /queryoutput
18/07/30 01:23:43 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your
iltin-java classes where applicable
Found 2 items
-rw-r--r--   1 acadgild supergroup          0 2018-07-30 01:16 /queryoutput/_SUCCESS
-rw-r--r--   1 acadgild supergroup         34 2018-07-30 01:16 /queryoutput/part-m-00000
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$
```

And then to check the result we use command :

hadoop fs -cat /queryoutput/part-m-00000

```
[acadgild@localhost ~]$ hadoop fs -cat /queryoutput/part-m-00000
18/07/30 01:27:08 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... usi
iltin-java classes where applicable
3,Rizvi,Syed,Whitefield,Bangalore
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$
```