

ASSIGNMENT 6.1

Task 1 : Export data present in sqoopout.

Input:-

```
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ hdfs dfs -cat /sqoopout/part-m-00000
18/02/13 04:29:35 WARN util.NativeCodeLoader: Unable to load native-hadoop libra
1,Shyam,Ram,Patna,Bihar
```

Command: `sqoop export --connect jdbc:mysql://localhost/simplidb --table Person1 --username root -P --export-dir /sqoopout.`

```
acadgild@localhost:~
1,Shyam,Ram,Patna,Bihar
[acadgild@localhost ~]$ sqoop export --connect jdbc:mysql://localhost/simplidb --table Person1 --username root -P --export-dir /sqoopout
Warning: /home/acadgild/install/sqoop/sqoop-1.4.6.bin_hadoop-2.0.4-alpha/./hcatalog does not exist! HCatalog jobs will fail.
Please set $HCAT_HOME to the root of your HCatalog installation.
Warning: /home/acadgild/install/sqoop/sqoop-1.4.6.bin_hadoop-2.0.4-alpha/./accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
18/02/13 04:34:05 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6
Enter password:
18/02/13 04:34:09 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
18/02/13 04:34:09 INFO tool.CodeGenTool: Beginning code generation
Tue Feb 13 04:34:09 IST 2018 WARN: Establishing SSL connection without server's identity verification is not recommended. According to My
requirements SSL connection must be established by default if explicit option isn't set. For compliance with existing applications not v
icate property is set to 'false'. You need either to explicitly disable SSL by setting useSSL=false, or set useSSL=true and provide trust
rification.
18/02/13 04:34:10 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM 'Person1' AS t LIMIT 1
18/02/13 04:34:10 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM 'Person1' AS t LIMIT 1
18/02/13 04:34:10 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /home/acadgild/install/hadoop/hadoop-2.6.5
Note: /tmp/sqoop-acadgild/compile/3f143efffa24a2b47ac16968ea030ab6/Person1.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
18/02/13 04:34:12 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-acadgild/compile/3f143efffa24a2b47ac16968ea030ab6/Person1.jar
18/02/13 04:34:12 INFO mapreduce.ExportJobBase: Beginning export of Person1
```

```

18/02/13 04:34:41 INFO mapreduce.Job: map 100% reduce 0%
18/02/13 04:34:42 INFO mapreduce.Job: Job job_1518445375024_0042 completed successfully
18/02/13 04:34:43 INFO mapreduce.Job: Counters: 30
  File System Counters
    FILE: Number of bytes read=0
    FILE: Number of bytes written=510176
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=568
    HDFS: Number of bytes written=0
    HDFS: Number of read operations=16
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=0
  Job Counters
    Launched map tasks=4
    Data-local map tasks=4
    Total time spent by all maps in occupied slots (ms)=58629
    Total time spent by all reduces in occupied slots (ms)=0
    Total time spent by all map tasks (ms)=58629
    Total vcore-milliseconds taken by all map tasks=58629
    Total megabyte-milliseconds taken by all map tasks=60036096
  Map-Reduce Framework
    Map input records=1
    Map output records=1
    Input split bytes=496
    Spilled Records=0
    Failed Shuffles=0
    Merged Map outputs=0
    GC time elapsed (ms)=601
    CPU time spent (ms)=7490
    Physical memory (bytes) snapshot=642973696
    Virtual memory (bytes) snapshot=8344297472
    Total committed heap usage (bytes)=454557696
  File Input Format Counters
    Bytes Read=0
  File Output Format Counters
    Bytes Written=0
18/02/13 04:34:43 INFO mapreduce.ExportJobBase: Transferred 568 bytes in 29.3372 seconds (19.3611 bytes/sec)
18/02/13 04:34:43 INFO mapreduce.ExportJobBase: Exported 1 records.

```

Output:- we have created the Person1 in simplidb database similar table like Person to get the output in the table.

```

acadgild@localhost:~
mysql> select * from Person1;
+-----+-----+-----+-----+-----+
| person_id | lname | fname | area | city |
+-----+-----+-----+-----+-----+
|          1 | Shyam | Ram   | Patna | Bihar |
+-----+-----+-----+-----+-----+
1 row in set (0.00 sec)

```

Input :-

acadgild@localhost:~

```
[acadgild@localhost ~]$ hdfs dfs -cat /queryoutput/part-m-00000
18/02/13 04:46:38 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform.
1,Shyam,Ram,Patna,Bihar
2,Tanya,Priya,Whitefiled,Bangalore
3,James,Brown,New York,United States
4,Jhon,Miller, Los Angeles ,United States
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ hdfs dfs -cat /queryoutput/part-m-00003
18/02/13 04:46:44 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform.
789,a,b,c,d
[acadgild@localhost ~]$
```

Command: `sqoop export --connect jdbc:mysql://localhost/simplidb --table Person2 --username root -P --export-dir /queryoutput`

```
[acadgild@localhost ~]$ sqoop export --connect jdbc:mysql://localhost/simplidb --table Person2 --username root -P --export-dir /queryoutput
Warning: /home/acadgild/install/sqoop/sqoop-1.4.6/bin/_hadoop-2.0.4-alpha/./hcatalog does not exist! HCatalog jobs will fail.
Please set $HCAT_HOME to the root of your HCatalog installation.
Warning: /home/acadgild/install/sqoop/sqoop-1.4.6/bin/_hadoop-2.0.4-alpha/./accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
18/02/13 04:42:13 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6
Enter password:
18/02/13 04:42:17 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
18/02/13 04:42:17 INFO tool.CodeGenTool: Beginning code generation
Tue Feb 13 04:42:17 IST 2018 WARN: Establishing SSL connection without server's identity verification is not recommended. According to MySQL
requirements SSL connection must be established by default if explicit option isn't set. For compliance with existing applications not using
icate property is set to 'false'. You need either to explicitly disable SSL by setting useSSL=false, or set useSSL=true and provide truststoc
rification.
18/02/13 04:42:18 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `Person2` AS t LIMIT 1
18/02/13 04:42:18 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `Person2` AS t LIMIT 1
18/02/13 04:42:18 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /home/acadgild/install/hadoop/hadoop-2.6.5
Note: /tmp/sqoop-acadgild/compile/2a24d917bb5dd18125df7729c6997638/Person2.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
18/02/13 04:42:20 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-acadgild/compile/2a24d917bb5dd18125df7729c6997638/Person2.jar
18/02/13 04:42:20 INFO mapreduce.ExportJobBase: Beginning export of Person2
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/common/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl
SLF4J: Found binding in [jar:file:/home/acadgild/install/hbase-1.2.6/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.cl
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
18/02/13 04:42:21 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where app
18/02/13 04:42:21 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar
18/02/13 04:42:22 INFO Configuration.deprecation: mapred.reduce.tasks.speculative.execution is deprecated. Instead, use mapreduce.reduce.spe
18/02/13 04:42:22 INFO Configuration.deprecation: mapred.map.tasks.speculative.execution is deprecated. Instead, use mapreduce.map.speculati
18/02/13 04:42:22 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps
18/02/13 04:42:22 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
```

```

acadgild@localhost:~
18/02/13 04:42:51 INFO mapreduce.Job: map 100% reduce 0%
18/02/13 04:42:51 INFO mapreduce.Job: Job job_1518445375024_0043 completed successfully
18/02/13 04:42:51 INFO mapreduce.Job: Counters: 31
  File System Counters
    FILE: Number of bytes read=0
    FILE: Number of bytes written=510188
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=1059
    HDFS: Number of bytes written=0
    HDFS: Number of read operations=25
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=0
  Job Counters
    Launched map tasks=4
    Other local map tasks=1
    Data-local map tasks=3
    Total time spent by all maps in occupied slots (ms)=58799
    Total time spent by all reduces in occupied slots (ms)=0
    Total time spent by all map tasks (ms)=58799
    Total vcore-milliseconds taken by all map tasks=58799
    Total megabyte-milliseconds taken by all map tasks=60210176
  Map-Reduce Framework
    Map input records=5
    Map output records=5
    Input split bytes=697
    Spilled Records=0
    Failed Shuffles=0
    Merged Map outputs=0
    GC time elapsed (ms)=686
    CPU time spent (ms)=7610
    Physical memory (bytes) snapshot=640765952
    Virtual memory (bytes) snapshot=8346451968
    Total committed heap usage (bytes)=452460544
  File Input Format Counters
    Bytes Read=0
  File Output Format Counters
    Bytes Written=0
18/02/13 04:42:51 INFO mapreduce.ExportJobBase: Transferred 1.0342 KB in 29.5671 seconds (35.8168 bytes/sec)
18/02/13 04:42:51 INFO mapreduce.ExportJobBase: Exported 5 records.
You have new mail in /var/spool/mail/acadgild

```

Output:-

```

acadgild@localhost:~
mysql> select * from Person2;
+-----+-----+-----+-----+-----+
| person_id | lname | fname | area | city |
+-----+-----+-----+-----+-----+
| 1 | Shyam | Ram | Patna | Bihar |
| 2 | Tanya | Priya | Whitefiled | Bangalore |
| 3 | James | Brown | New York | United States |
| 4 | Jhon | Miller | Los Angeles | United States |
| 789 | a | b | c | d |
+-----+-----+-----+-----+-----+
5 rows in set (0.00 sec)

```

Task 2: Export data present in sqoopout with parameter person_id = 3.

Update or insert is not supported in sqoop export execution for MySQL databases.

