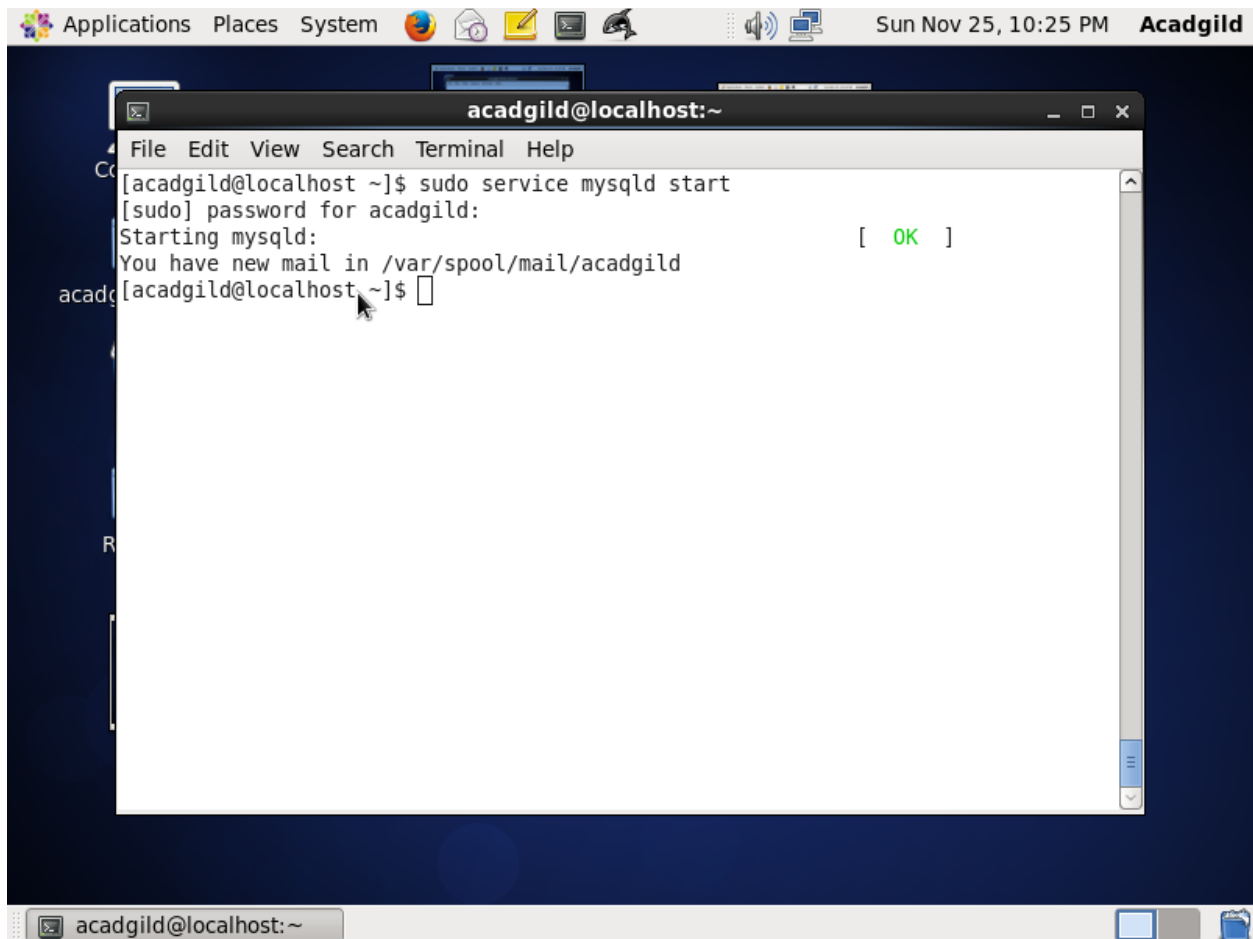


Sqoop Assignment

Task 1

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

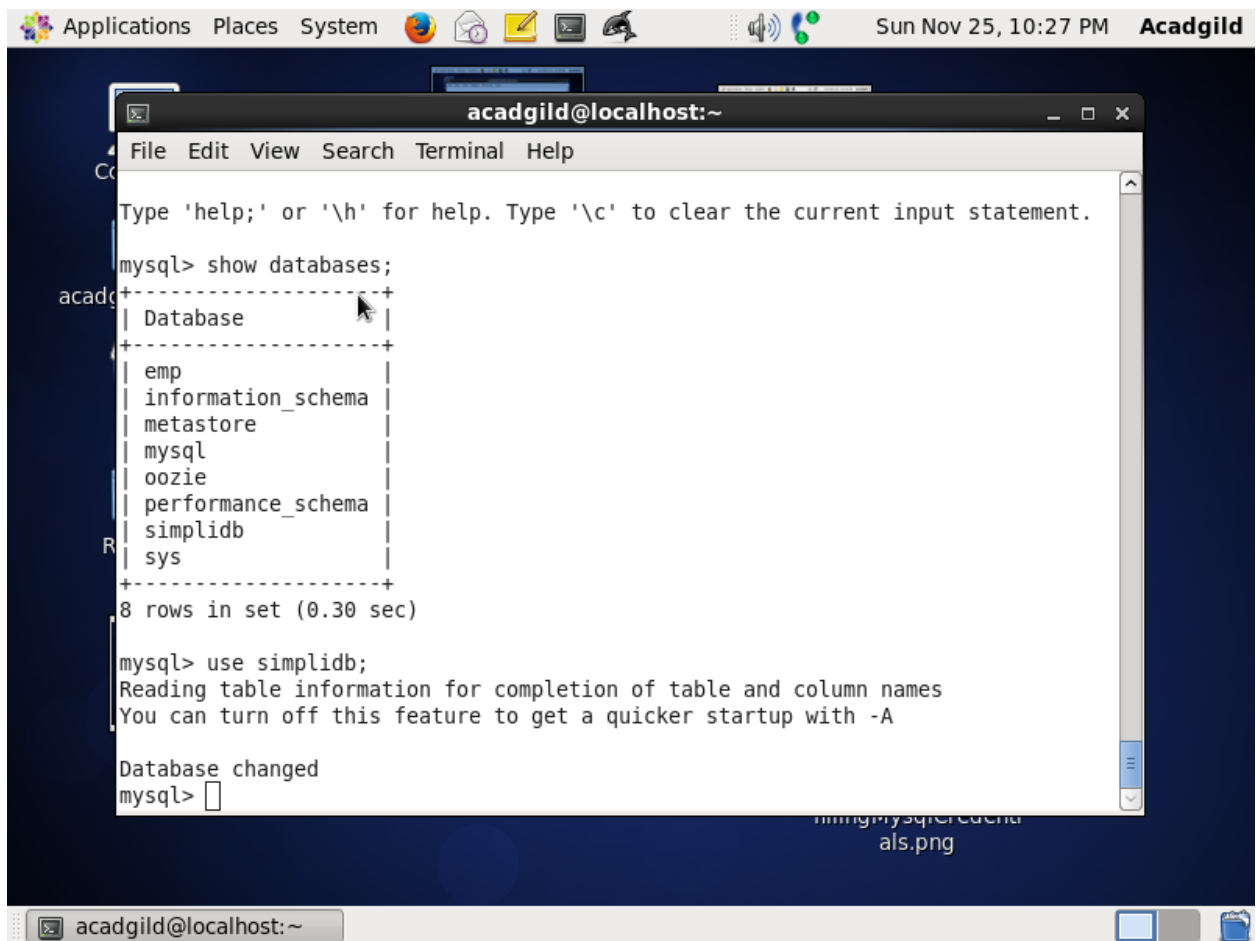
Solution: Starting the mysql services:



The screenshot shows a Linux desktop environment with a terminal window open. The terminal title is "acadgild@localhost:~". The terminal output shows the command "sudo service mysqld start" being executed. The prompt "[sudo] password for acadgild:" is shown, followed by "Starting mysqld:" and a green "[OK]" status. A message "You have new mail in /var/spool/mail/acadgild" is also displayed. The terminal prompt is now "[acadgild@localhost ~]\$". The desktop background is dark blue, and the top panel shows the date "Sun Nov 25, 10:25 PM" and the username "Acadgild".

```
acadgild@localhost:~$ sudo service mysqld start
[sudo] password for acadgild:
Starting mysqld: [ OK ]
You have new mail in /var/spool/mail/acadgild
acadgild@localhost ~]$
```

Showing the databases and using the required database:



The screenshot shows a terminal window titled 'acadgild@localhost:~' with a menu bar (File, Edit, View, Search, Terminal, Help). The terminal displays the following text:

```
Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql> show databases;
+-----+
| Database |
+-----+
| emp      |
| information_schema |
| metastore |
| mysql    |
| oozie     |
| performance_schema |
| simplidb  |
| sys      |
+-----+
8 rows in set (0.30 sec)

mysql> use simplidb;
Reading table information for completion of table and column names
You can turn off this feature to get a quicker startup with -A

Database changed
mysql> 
```

The terminal window is part of a desktop environment with a taskbar at the top showing icons for Applications, Places, System, and network status. The system clock indicates 'Sun Nov 25, 10:27 PM' and the username 'Acadgild' is in the top right corner.

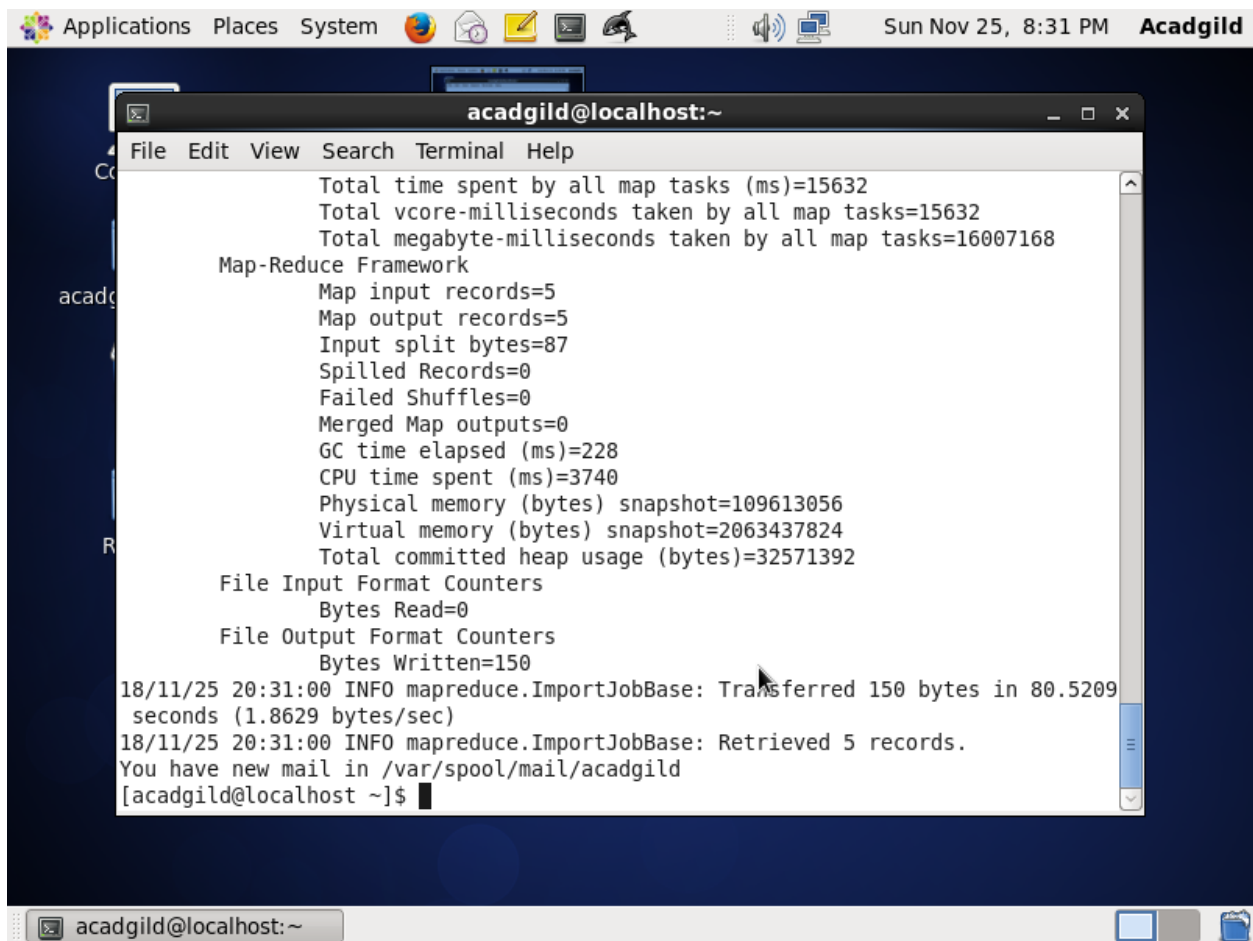
Showing the records in the Person Table:

Applications Places System Sun Nov 25, 10:28 PM Acadgild

```
acadm@localhost:~  
File Edit View Search Terminal Help  
Database changed  
mysql> show Tables  
-> ;  
+-----+  
| Tables_in_simplidb |  
+-----+  
| Person              |  
+-----+  
1 row in set (0.00 sec)  
  
mysql> SELECT * FROM Person;  
+-----+  
| 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)  
  
mysql> 
```

Importing the records of the table in HDFS:

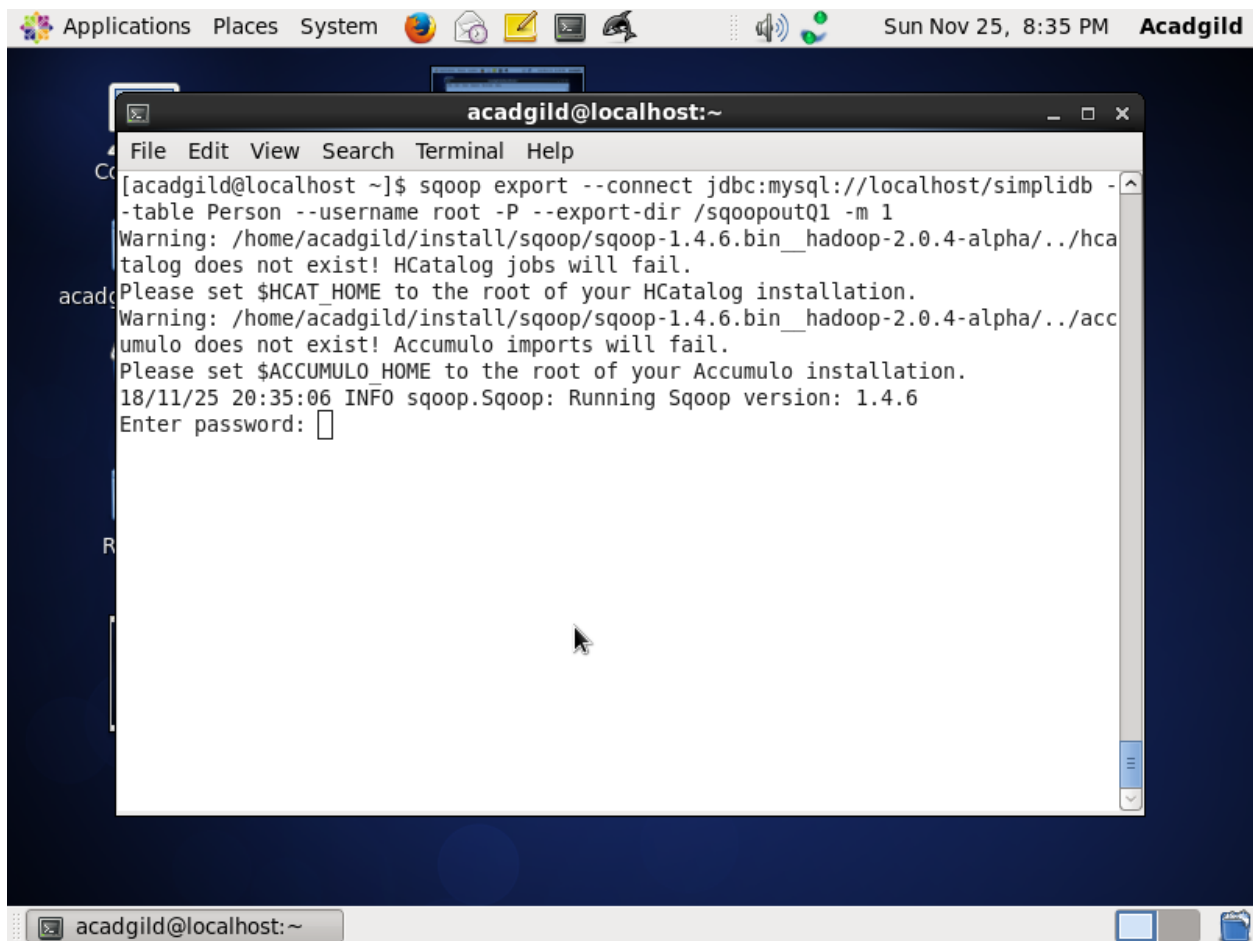
```
acadm@localhost:~$ sqoop import --connect jdbc:mysql://localhost/simplidb -
--table Person --username root -P --target-dir /sqoopoutQ1 -m 1
Warning: /home/acadm/install/sqoop/sqoop-1.4.6.bin__hadoop-2.0.4-alpha/./hca
atalog does not exist! HCatalog jobs will fail.
Please set $HCAT HOME to the root of your HCatalog installation.
Warning: /home/acadm/install/sqoop/sqoop-1.4.6.bin__hadoop-2.0.4-alpha/./acc
umulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO HOME to the root of your Accumulo installation.
18/11/25 20:29:16 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6
Enter password:
18/11/25 20:29:22 INFO manager.MySQLManager: Preparing to use a MySQL streaming
resultset.
18/11/25 20:29:22 INFO tool.CodeGenTool: Beginning code generation
Sun Nov 25 20:29:23 IST 2018 WARN: Establishing SSL connection without server's
identity verification is not recommended. According to MySQL 5.5.45+, 5.6.26+ an
d 5.7.6+ requirements SSL connection must be established by default if explicit
option isn't set. For compliance with existing applications not using SSL the ve
rifyServerCertificate property is set to 'false'. You need either to explicitly
disable SSL by setting useSSL=false, or set useSSL=true and provide truststore f
or server certificate verification.
18/11/25 20:29:26 INFO manager.SqlManager: Executing SQL statement: SELECT t.* F
ROM `Person` AS t LIMIT 1
18/11/25 20:29:26 INFO manager.SqlManager: Executing SQL statement: SELECT t.* F
```



The screenshot shows a Linux desktop environment with a terminal window titled 'acadgild@localhost:~'. The terminal displays the output of a Hadoop MapReduce job. The output includes summary statistics for all map tasks, detailed metrics for the Map-Reduce Framework, and file input/output counters. It also shows log messages from the mapreduce.ImportJobBase class and a system message about new mail.

```
acacgild@localhost:~  
File Edit View Search Terminal Help  
Total time spent by all map tasks (ms)=15632  
Total vcore-milliseconds taken by all map tasks=15632  
Total megabyte-milliseconds taken by all map tasks=16007168  
Map-Reduce Framework  
  Map input records=5  
  Map output records=5  
  Input split bytes=87  
  Spilled Records=0  
  Failed Shuffles=0  
  Merged Map outputs=0  
  GC time elapsed (ms)=228  
  CPU time spent (ms)=3740  
  Physical memory (bytes) snapshot=109613056  
  Virtual memory (bytes) snapshot=2063437824  
  Total committed heap usage (bytes)=32571392  
File Input Format Counters  
  Bytes Read=0  
File Output Format Counters  
  Bytes Written=150  
18/11/25 20:31:00 INFO mapreduce.ImportJobBase: Transferred 150 bytes in 80.5209  
seconds (1.8629 bytes/sec)  
18/11/25 20:31:00 INFO mapreduce.ImportJobBase: Retrieved 5 records.  
You have new mail in /var/spool/mail/acadgild  
[acadgild@localhost ~]$
```

Exporting the records from HDFS to mysql:



The screenshot shows a Linux desktop environment with a terminal window titled "acadgild@localhost:~". The terminal displays the following commands and output:

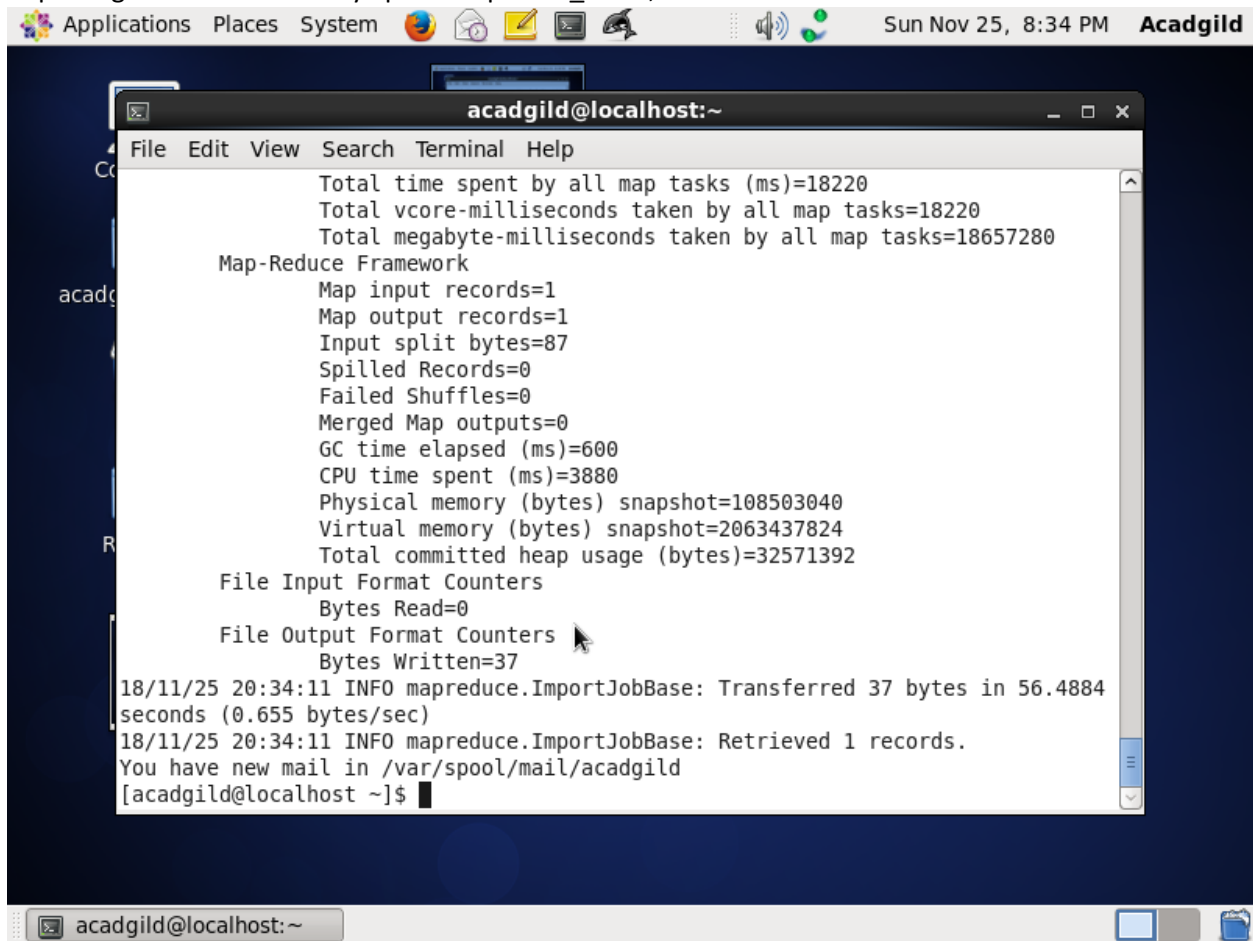
```
[acadgild@localhost ~]$ sqoop export --connect jdbc:mysql://localhost/simplidb -  
-table Person --username root -P --export-dir /sqoopoutQ1 -m 1  
Warning: /home/acadgild/install/sqoop/sqoop-1.4.6.bin__hadoop-2.0.4-alpha/./hca  
talog does not exist! HCatalog jobs will fail.  
Please set $HCAT_HOME to the root of your HCatalog installation.  
Warning: /home/acadgild/install/sqoop/sqoop-1.4.6.bin__hadoop-2.0.4-alpha/./acc  
umulo does not exist! Accumulo imports will fail.  
Please set $ACCUMULO_HOME to the root of your Accumulo installation.  
18/11/25 20:35:06 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6  
Enter password: [ ]
```

Task2

Use Sqoop tool to export data present in SQOOPOUT folder made while demo of Import table with parameter person_id =3.

Solution:

Importing the record from mysql where person_id = 3, to hdfs.



The screenshot shows a Linux desktop environment with a terminal window titled 'acadgild@localhost:~'. The terminal displays the output of a Hadoop MapReduce job. The statistics are as follows:

```
File Edit View Search Terminal Help
Total time spent by all map tasks (ms)=18220
Total vcore-milliseconds taken by all map tasks=18220
Total megabyte-milliseconds taken by all map tasks=18657280
Map-Reduce Framework
  Map input records=1
  Map output records=1
  Input split bytes=87
  Spilled Records=0
  Failed Shuffles=0
  Merged Map outputs=0
  GC time elapsed (ms)=600
  CPU time spent (ms)=3880
  Physical memory (bytes) snapshot=108503040
  Virtual memory (bytes) snapshot=2063437824
  Total committed heap usage (bytes)=32571392
File Input Format Counters
  Bytes Read=0
File Output Format Counters
  Bytes Written=37
18/11/25 20:34:11 INFO mapreduce.ImportJobBase: Transferred 37 bytes in 56.4884
seconds (0.655 bytes/sec)
18/11/25 20:34:11 INFO mapreduce.ImportJobBase: Retrieved 1 records.
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$
```

The terminal window is part of a desktop environment with a taskbar at the top showing 'Applications', 'Places', 'System', and the date 'Sun Nov 25, 8:34 PM'. The user 'Acadgild' is logged in. The terminal window has a menu bar with 'File', 'Edit', 'View', 'Search', 'Terminal', and 'Help'. The status bar at the bottom shows 'acadgild@localhost:~' and some system icons.

Exporting the record from hdfs to mysql:

```
acadm@localhost:~  
File Edit View Search Terminal Help  
[acadm@localhost ~]$ sqoop export --connect jdbc:mysql://localhost/simplidb -  
-table Person --username root -P --export-dir /sqoopoutQ2 -m 1  
Warning: /home/acadm/install/sqoop/sqoop-1.4.6.bin__hadoop-2.0.4-alpha/./hca  
talog does not exist! HCatalog jobs will fail.  
Please set $HCAT_HOME to the root of your HCatalog installation.  
Warning: /home/acadm/install/sqoop/sqoop-1.4.6.bin__hadoop-2.0.4-alpha/./acc  
umulo does not exist! Accumulo imports will fail.  
Please set $ACCUMULO_HOME to the root of your Accumulo installation.  
18/11/25 22:03:08 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6  
Enter password: 
```



```
acadm@localhost:~  
File Edit View Search Terminal Help  
[acadm@localhost ~]$ sqoop import --connect jdbc:mysql://localhost/simplidb -  
-table Person --username root -P --target-dir /sqoopoutQ2 -m 1 --where person_id  
=3  
Warning: /home/acadm/install/sqoop/sqoop-1.4.6.bin__hadoop-2.0.4-alpha/./hca  
talog does not exist! HCatalog jobs will fail.  
Please set $HCAT HOME to the root of your HCatalog installation.  
Warning: /home/acadm/install/sqoop/sqoop-1.4.6.bin__hadoop-2.0.4-alpha/./acc  
umulo does not exist! Accumulo imports will fail.  
Please set $ACCUMULO HOME to the root of your Accumulo installation.  
18/11/25 20:32:44 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6  
Enter password: 
```