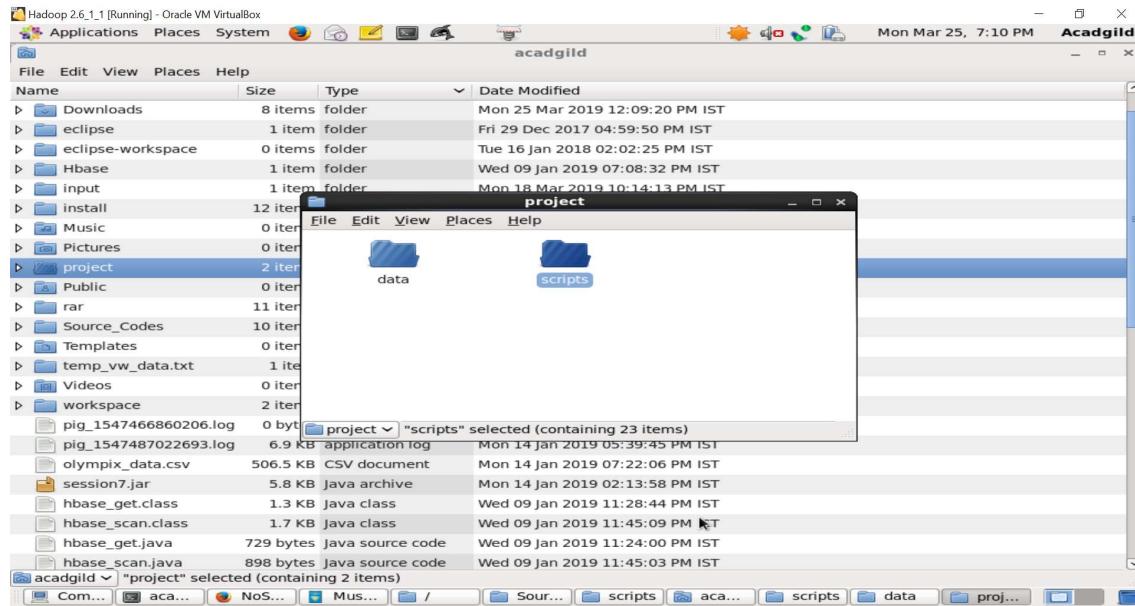


## Vijay's Final project

Based on the description provided and the same code, this project has been executed. Few changes were made to the data and to the scripts when errors were faced during execution. Details of those errors and the solution are given below.

- 1) Code library copied to Acadgild VM. A new folder Project was created under /home/acadgild and two new sub folders data & scripts created inside it.



- 2) All the sample scripts provided were copied over to the scripts folder and the permission for it was changed as given below.

```
[acadgild@localhost ~]$ chmod 774 /home/acadgild/project/scripts/*
```

- 3) The next step was to generate sample data for web and mob interfaces. For web, following Python program was used. This program was modified a little to use 2017 & 2018 dates instead of the original 2016 dates. The program was executed as given below

```
[acadgild@localhost ~]$ python /home/acadgild/project/scripts/generate_web_data.py
```

- 4) The output created by the program was file.xml. The file contents are given below.

```

<records>
  <record>
    <user_id>U119</user_id>
    <song_id>S208</song_id>
    <artist_id>A301</artist_id>
    <timestamp>2017-06-09 22:12:36</timestamp>
    <start_ts>2017-07-10 01:38:09</start_ts>
    <end_ts>2017-06-09 22:12:36</end_ts>
    <geo_cd>U</geo_cd>
    <station_id>ST406</station_id>
    <song_end_type>2</song_end_type>
    <like>1</like>
    <dislike>0</dislike>
  </record>
  <record>
    <user_id>U102</user_id>
    <song_id>S200</song_id>
    <artist_id>A303</artist_id>
    <timestamp>2018-05-09 08:09:22</timestamp>
    <start_ts>2017-07-10 01:38:09</start_ts>
    <end_ts>2017-06-09 22:12:36</end_ts>
    <geo_cd>E</geo_cd>
    <station_id>ST407</station_id>
    <song_end_type>1</song_end_type>
    <like>1</like>
  </record>

```

- 5) To generate mobile data, following python code was executed and the output is given below.

```
[acadgild@localhost ~]$ python /home/acadgild/project/scripts/generate_mob_data.py
```

```

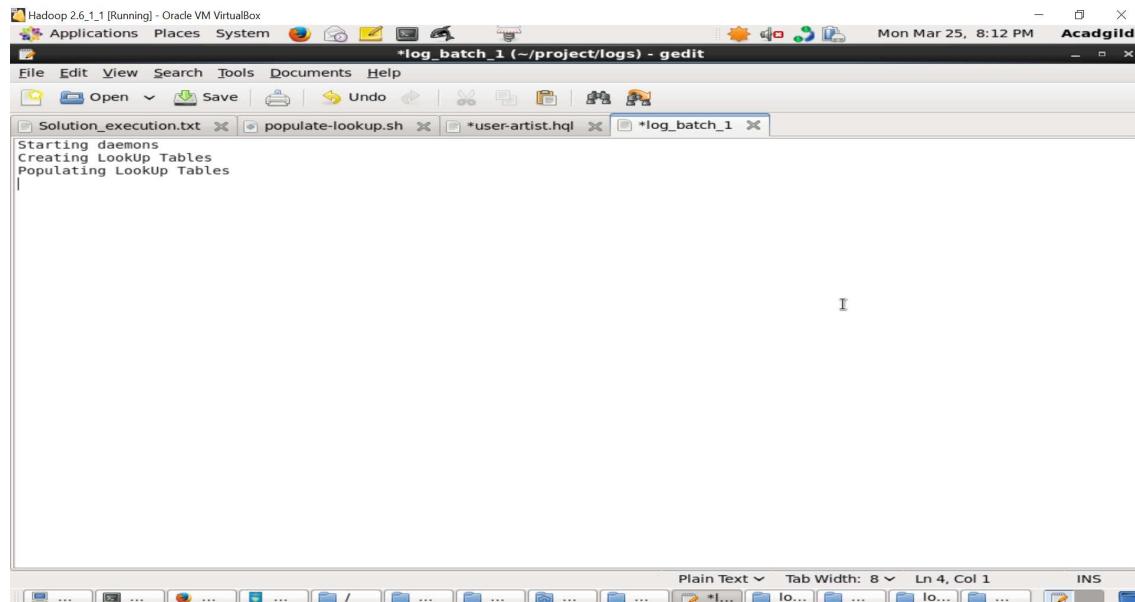
file.txt (-/project/data/mob) - gedit
File Edit View Search Tools Documents Help
File Open Save Undo Redo Cut Copy Paste Insert Find Replace Delete Select All
file.txt
U112,S209,A304,1465230523,1475130523,1465230523,E,ST400,2,0,1
U107,S204,A303,1465130523,1475130523,1485130523,AU,ST405,1,1,0
U107,S201,A304,1475130523,1475130523,1485130523,E,ST403,1,1,1
U104,S210,A301,1465130523,1465230523,1475130523,E,ST412,0,0,0
U113,S207,A301,1465230523,1465230523,1475130523,E,ST409,1,0,0
U201,S201,A302,1495130523,1485130523,1465230523,AU,ST401,3,0,0
U123,S204,A303,1465230523,1485130523,1465230523,AU,ST413,1,0,1
U119,S203,A305,1465230523,1485130523,1465130523,AU,ST412,2,0,1
U115,S207,A302,1465130523,1475130523,1465230523,ST400,1,1,0
U102,S203,,1465230523,1485130523,1465230523,U,ST408,0,1,1
U100,S201,A301,1495130523,1475130523,1465230523,AP,ST405,1,1,0
U105,S210,A304,1495130523,1485130523,1465230523,E,ST401,1,1,1
U119,S207,A301,1495130523,1465230523,1485130523,U,ST413,3,1,0
U117,S208,A303,1465230523,1485130523,1485130523,U,ST403,3,1,1
U104,S204,A301,1475130523,1465230523,1485130523,U,ST400,3,0,0
U107,S207,A301,1465130523,1465230523,1475130523,A,ST409,3,1,1
U101,S208,A304,1475130523,1465230523,1475130523,AP,ST411,1,0,0
U111,S204,A302,1475130523,1475130523,1465130523,AP,ST405,3,1,1
U114,S210,A304,1465130523,1485130523,1475130523,AU,ST410,3,1,1
U115,S207,A300,1465230523,1485130523,1475130523,E,ST412,2,1,0

```

- 6) Using the first shell script, a batch id was created and a new log file using that batch id was created. All the Hadoop daemons were started using start-all.sh. Since my VM had these services executing already, following messages were displayed. Hbase was started using start-hbase and the job history server was started using mr-jobhistory-daemon.sh

```
[acadgild@localhost ~]$ sh /home/acadgild/project/scripts/start-daemons.sh
This script is Deprecated. Instead use start-dfs.sh and start-yarn.sh
19/03/25 19:39:41 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Starting namenodes on [localhost]
localhost: namenode running as process 4205. Stop it first.
localhost: datanode running as process 4308. Stop it first.
Starting secondary namenodes [0.0.0.0]
0.0.0.0: secondarynamenode running as process 4526. Stop it first.
19/03/25 19:39:50 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
starting yarn daemons
resourcemanager running as process 4674. Stop it first.
localhost: nodemanager running as process 4777. Stop it first.
localhost: zookeeper running as process 23593. Stop it first.
master running as process 23688. Stop it first.
regionserver running as process 23782. Stop it first.
historyserver running as process 23883. Stop it first.
[acadgild@localhost ~]$
```

First few statements written to the log file log\_batch\_1 is given below



- 7) populate-lookup.sh -- Populates lookup tables

- Using the input files provided in the sample code package, the data was loaded into the lookup tables. Three lookup tables were created in Hbase and the data was loaded from the three input files provided.
- Once the data was loaded into Hbase, a new table users\_artists was created in Hive and the data from the local file user\_artist.txt was loaded into it.
- The initial load command failed while executing the shell script as the input file was missing. Hence the load into users\_artists was executed in hive shell directly. Once

the data was loaded, it was selected for verification.

```
hive> LOAD DATA LOCAL INPATH '/home/acadgild/project/lookupfiles/user-artist.txt' OVERWRITE INTO TABLE users_artists;
Loading data to table project.users_artists
OK
Time taken: 3.182 seconds
hive> select * from users_artists;
OK
U100    ["A300","A301","A302"]
U101    ["A301","A302"]
U102    ["A302"]
U103    ["A303","A301","A302"]
U104    ["A304","A301"]
U105    ["A305","A301","A302"]
U106    ["A301","A302"]
U107    ["A302"]
U108    ["A300","A303","A304"]
U109    ["A301","A303"]
U110    ["A302","A301"]
U111    ["A303","A301"]
U112    ["A304","A301"]
U113    ["A305","A302"]
U114    ["A300","A301","A302"]
Time taken: 3.746 seconds, Fetched: 15 row(s)
```

## 8) dataformatting.sh -- Performs data formatting

- The formatting of the data as per the project requirements was performed by executing the shell script dataformatting.sh
- This script, first removes the existing paths in Hadoop for mobile and web. Once the folders are removed from Hadoop, new folders are created and the content from the local mob & web are copied into it.
- Before executing the pig code, piggybank.jar was copied to /home/acadgild/project/lib/ library in local.
- The input file was loaded into Pig using the XMLLoader option as the web data was generated in XML format as shown above.
- This pig program converted the web data into the format similar to mobile data. The timestamp field was converted from YYYY-MM-DD HH:MM:SS to UNIX timestamp long integers
- The converted output was stored in a comma separated format inside the formattedweb folder

```
2019-03-25 21:27:09,258 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost/127.0.1:8032
2019-03-25 21:27:09,287 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job history server
2019-03-25 21:27:09,497 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost/127.0.1:8032
2019-03-25 21:27:09,535 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job history server
2019-03-25 21:27:09,734 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at localhost/127.0.1:8032
2019-03-25 21:27:09,772 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job history server
2019-03-25 21:27:09,953 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Success !
2019-03-25 21:27:10,063 [main] INFO org.apache.pig.Main - Pig script completed in 1 minute and 688 milliseconds (60688 ms)
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/acadgild/install/hive/apache-hive-2.3.2-bin/lib/log4j-slf4j-impl-2.6.2.jar!/:org/slf4j/impl/StaticLoggerBinder.class]
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/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]

Logging initialized using configuration in jar:file:/home/acadgild/install/hive/apache-hive-2.3.2-bin/lib/hive-common-2.3.2.jar!/hive-log4j2.properties Async: true
OK
Time taken: 16.739 seconds
OK
Time taken: 0.63 seconds
Loading data to table project.formatted_input partition (batchid=1)
OK
Time taken: 5.028 seconds
Loading data to table project.formatted_input partition (batchid=1)
OK
Time taken: 2.345 seconds
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$
```

- A new table formatted\_input created in Hive to store the formatted output from the Pig program.
- Once the table is created, both the formatted web data and the generated mobile data was loaded into the table. Sample select from the table is given below.

```

Hadoop 2.6.1_1 [Running] - Oracle VM VirtualBox
Applications Places System Mon Mar 25, 9:39 PM Acadgild
acadgild@localhost:~ %
File Edit View Search Terminal Help
hive> select * from formatted_input;
OK
U112      S209      A304      1465230523      1475130523      1465230523      E      ST400      2      0      1      1
U107      S204      A303      1465130523      1475130523      1485130523      AU     ST405      1      1      0      1
U107      S201      A304      1475130523      1475130523      1485130523      E      ST403      1      1      1      1
U104      S210      A301      1465130523      1465230523      1475130523      E      ST412      0      0      0      1
U113      S207      A301      1465230523      1465230523      1475130523      E      ST409      1      0      0      1
U120      S204      A303      1465230523      1465230523      1485130523      AU     ST401      3      0      0      1
U119      S203      A305      1465230523      1485130523      1465130523      AU     ST415      2      0      1      1
U115      S207      A302      1465130523      1475130523      1465230523      ST400      1      1      0      1
U102      S203      A302      1465230523      1485130523      1465230523      U      ST408      0      1      1      1
U105      S210      A304      1495130523      1475130523      1465230523      AP     ST405      1      1      0      1
U119      S207      A301      1495130523      1465230523      1485130523      U      ST413      3      1      0      1
U117      S208      A303      1465230523      1485130523      1485130523      U      ST403      3      1      1      1
U104      S204      A301      1475130523      1465230523      1465130523      U      ST400      3      0      0      1
U107      S207      A301      1475130523      1465230523      1475130523      AP     ST409      3      1      1      1
U101      S208      A304      1475130523      1465130523      1475130523      AP     ST411      1      0      0      1
U111      S204      A302      1465130523      1475130523      1465230523      AP     ST415      3      1      1      1
U114      S210      A304      1465130523      1485130523      1475130523      AU     ST410      3      1      1      1
U115      S207      A300      1465230523      1485130523      1475130523      E      ST412      2      1      0      1
U119      S208      A301      1497026556      1499630889      1497026556      U      ST406      2      1      0      1
U102      S200      A303      1525833562      1499630889      1497026556      E      ST407      1      1      1      1
U115      S209      A301      1494399262      1497026556      1525833562      E      ST404      2      0      1      1
U117      S202      A301      1499630889      1497026556      1494399262      AU     ST402      2      1      0      1
U117      S208      A304      1525833562      1499630889      1499630889      A      ST409      0      1      1      1
S203      A301      1499630889      1494399262      1494399262      A      ST410      2      0      1      1
U114      S203      A305      1499630889      1497026556      1525833562      AP     ST405      3      0      1      1
U103      S208      A300      1499630889      1494399262      1525833562      AP     ST405      0      0      1      1
U102      S202      A300      1525833562      1499630889      1494399262      AP     ST406      1      0      1      1
U117      S200      A300      1497026556      1499630889      1494399262      AU     ST410      0      0      0      1
U100      S209      A301      1497026556      1499630889      1497026556      U      ST406      2      1      1      1
U114      S205      A300      1494399262      1497026556      1525833562      E      ST414      3      1      0      1
U113      S203      A304      1525833562      1499630889      1496630889      AU     ST415      2      0      0      1
U113      S210      A302      1497026556      1525833562      1494399262      E      ST403      1      0      1      1
U100      S202      A300      1497026556      1499630889      1494399262      A      ST401      2      1      0      1
U107      S204      A303      1494399262      1497026556      1497026556      AP     ST407      1      0      1      1
U117      S208      A301      1499630889      1525833562      1494399262      E      ST411      1      1      0      1

```

9) `hive -f /home/acadgild/project/scripts/create_hive_hbase_lookup.hql`

- To perform the data enrichment, three external tables were created in Hive pointing to the Hbase lookup tables created earlier.
- The link between Hive and HBase was established using the HBaseStorageHandler option and the Hbase table name and the table schema was provided

```

Hadoop 2.6.1_1 [Running] - Oracle VM VirtualBox
Applications Places System Mon Mar 25, 9:42 PM Acadgild
acadgild@localhost:~ %
File Edit View Search Terminal Help
[acadgild@localhost ~]$ hive -f /home/acadgild/project/scripts/create_hive_hbase_lookup.hql
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/acadgild/install/hive/apache-hive-2.3.2-bin/lib/log4j-slf4j-impl-2.6.2.jar!/org/slf4j
    /1.7.25/log4jLogger.class]
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/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]

Logging initialized using configuration in jar:file:/home/acadgild/install/hive/apache-hive-2.3.2-bin/lib/hive-common-2.3.2.j
ar!/hive-log4j2.properties Async: true
OK
Time taken: 15.762 seconds
OK
Time taken: 5.879 seconds
OK
Time taken: 0.537 seconds
OK
Time taken: 0.615 seconds
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ █

```

c. The data in the new Hive lookup tables are given below.

```
Hadoop 2.6.1 [Running] - Oracle VM VirtualBox
Applications Places System Mon Mar 25, 9:45 PM Acadgild
File Edit View Search Terminal Help
acadgild@localhost:~>

hive> show tables;
OK
formatted_input
song_artist_map
station_geo_map
subscribed_users
users_artists
Time taken: 0.429 seconds, Fetched: 5 row(s)
hive> select * from song_artist_map;
OK
S200    A300
S201    A301
S202    A302
S203    A303
S204    A304
S205    A301
S206    A302
S207    A303
S208    A304
S209    A305
Time taken: 5.246 seconds, Fetched: 10 row(s)
hive> select * from station_geo_map;
OK
ST400    A
ST401    AU
ST402    AP
ST403    J
ST404    E
ST405    A
ST406    AU
ST407    AP
ST408    E
ST409    A
ST410    A
ST411    A
ST412    AP
ST413    J
ST414    E
Time taken: 1.107 seconds, Fetched: 15 row(s)

Hadoop 2.6.1 [Running] - Oracle VM VirtualBox
Applications Places System Mon Mar 25, 9:45 PM Acadgild
File Edit View Search Terminal Help
acadgild@localhost:~>

hive> select * from station_geo_map;
OK
ST400    A
ST401    AU
ST402    AP
ST403    J
ST404    E
ST405    A
ST406    AU
ST407    AP
ST408    E
ST409    E
ST410    A
ST411    A
ST412    AP
ST413    J
ST414    E
Time taken: 5.246 seconds, Fetched: 10 row(s)
hive> select * from subscribed_users;
OK
U100    1465230523    1465130523
U101    1465230523    1475130523
U102    1465230523    1475130523
U103    1465230523    1475130523
U104    1465230523    1475130523
U105    1465230523    1475130523
U106    1465230523    1485130523
U107    1465230523    1455130523
U108    1465230523    1465230623
U109    1465230523    1475130523
U110    1465230523    1475130523
U111    1465230523    1475130523
U112    1465230523    1475130523
U113    1465230523    1465130523
U114    1465230523    1468130523
Time taken: 1.18 seconds, Fetched: 15 row(s)
hive>
```

## 10) data\_enrichment.sh -- Performs data enrichment and cleaning

- a. As per the project requirements, the missing fields in the input data is enriched using the look up tables created above
- b. New Hive table enriched\_data is created. The timestamp column was renamed as time\_stamp and like column was renamed to like1.
- c. The rows that are missing column values, the data is enriched by looking up the lookup tables and then loaded into this enriched\_data table.

```
[acadgild@localhost ~]$ sh /home/acadgild/project/scripts/data_enrichment.sh
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/acadgild/install/hive/apache-hive-2.3.2-bin/lib/log4j-slf4j-impl-2.6.2.jar!/org/slf4j/impl/StaticLoggerBinder.class]
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/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]

Logging initialized using configuration in jar:file:/home/acadgild/install/hive/apache-hive-2.3.2-bin/lib/hive-common-2.3.2.jar!/hive-log4j2.properties Async: true
OK
Time taken: 17.921 seconds
OK
Time taken: 0.998 seconds
No Stats for project@formatted_input, Columns: start_ts, song_id, time_stamp, user_id, end_ts, dislike, station_id, geo_cd, like1, song_end_type
No Stats for project@station_geo_map, Columns: station_id, geo_cd
No Stats for project@song_artist_map, Columns: song_id, artist_id
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X releases.
Query ID = acadgild_20190325215617_b1f02elf-4ad4-4c9b-b49f-58c5711dc5a3
Total jobs = 2
Launching Job 1 out of 2
Number of reduce tasks not specified. Estimated from input data size: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1553492063162_0003, Tracking URL = http://localhost:8088/proxy/application_1553492063162_0003/
Kill Command = /home/acadgild/install/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1553492063162_0003
Hadoop job information for Stage-1: number of mappers: 3; number of reducers: 1
2019-03-25 21:57:03,795 Stage-1 map = 0%, reduce = 0%
2019-03-25 21:57:48,205 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 5.66 sec
2019-03-25 21:57:51,841 Stage-1 map = 100%, reduce = 0% Cumulative CPU 11.1 sec
2019-03-25 21:58:05,570 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 14.67 sec
MapReduce Total cumulative CPU time: 14 seconds 670 msec
Ended Job = job_1553492063162_0003
Launching Job 2 out of 2
Number of reduce tasks not specified. Estimated from input data size: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1553492063162_0004, Tracking URL = http://localhost:8088/proxy/application_1553492063162_0004/
Kill Command = /home/acadgild/install/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1553492063162_0004
Hadoop job information for Stage-2: number of mappers: 2; number of reducers: 1
2019-03-25 21:58:31,169 Stage-2 map = 0%, reduce = 0%
2019-03-25 21:58:46,069 Stage-2 map = 50%, reduce = 0%, Cumulative CPU 2.54 sec
2019-03-25 21:58:57,815 Stage-2 map = 100%, reduce = 0%, Cumulative CPU 7.16 sec
2019-03-25 21:59:12,306 Stage-2 map = 100%, reduce = 100%, Cumulative CPU 12.7 sec
MapReduce Total cumulative CPU time: 12 seconds 700 msec
Ended Job = job_1553492063162_0004
Loading data to table project.enriched_data partition (batchid=null, status=null)

Loaded : 2/2 partitions.
  Time taken to load dynamic partitions: 2.663 seconds
  Time taken for adding to write entity : 0.014 seconds
MapReduce Jobs Launched:
Stage-Stage-1: Map: 3 Reduce: 1 Cumulative CPU: 14.67 sec HDFS Read: 49538 HDFS Write: 3088 SUCCESS
Stage-Stage-2: Map: 2 Reduce: 1 Cumulative CPU: 12.7 sec HDFS Read: 24165 HDFS Write: 3239 SUCCESS
Total MapReduce CPU Time Spent: 27 seconds 370 msec
OK
Time taken: 180.957 seconds
19/03/25 21:59:21 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
19/03/25 21:59:26 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$
```

d. Data loaded in to this enriched table is selected

```

hive> select * from enriched_data;
OK
U102      S200      A300      1525833562      1499639889      1497826556      AP      ST407      1      1      1      1      1      fail
U107      S201      A301      1475130523      1485130523      J      ST403      1      1      1      1      1      fail
S201      A301      1495130523      1485130523      1465230523      AU      ST401      3      0      0      1      1      fail
U102      S202      A302      1525833562      1497826556      AU      ST401      0      1      0      1      1      fail
U119      S203      A303      1465230523      1485130523      NULL      ST415      2      0      1      1      1      fail
S203      A303      1499639889      1494399262      A      ST410      2      0      1      1      1      fail
U113      S203      A303      1525833562      1499639889      NULL      ST415      2      0      0      1      1      fail
U102      S203      A303      1465230523      1485130523      1465230523      E      ST408      0      1      1      1      1      fail
U111      S204      A304      1475130523      1485130523      A      ST405      3      1      1      1      1      fail
U117      S206      A302      1497026556      1499630889      1497026556      E      ST408      1      1      1      1      1      fail
U117      S207      A303      1465130523      1475130523      1465230523      A      ST400      1      1      0      1      1      fail
U107      S207      A303      1465130523      1485130523      1475130523      E      ST409      3      1      1      1      1      fail
U117      S208      A304      1499630889      1499630889      1499630889      J      ST409      3      1      1      1      1      fail
U117      S208      A304      1525833562      1499630889      E      ST409      0      1      1      1      1      fail
U100      S209      A305      1497026556      1499630889      1497026556      AU      ST406      2      1      1      1      1      fail
U105      S210      NULL      1495130523      1485130523      1465230523      AU      ST401      1      1      1      1      1      fail
U113      S210      NULL      1497026556      1525833562      1494399262      J      ST403      1      0      1      1      1      fail
U114      S210      NULL      1465130523      1485130523      1475130523      A      ST410      3      1      1      1      1      fail
U104      S210      NULL      1465130523      1465230523      1475130523      AP      ST412      0      0      0      1      1      fail
U117      S210      NULL      1497026556      1499630889      1494399262      A      ST410      0      0      0      1      1      fail
U100      S201      A301      1495130523      1475130523      1465230523      A      ST405      1      1      0      1      1      pass
U100      S202      A302      1497026556      1499630889      1494399262      AU      ST401      2      1      0      1      1      pass
U117      S202      A302      1499630889      1497026556      1494399262      AP      ST402      2      1      0      1      1      pass
U116      S202      A302      1499630889      1497026556      1497826556      J      ST403      0      1      1      1      1      pass
U111      S202      A302      1499630889      1497026556      1497826556      AP      ST402      2      1      0      1      1      pass
U114      S203      A303      1499630889      1497026556      1525833562      AU      ST406      1      0      1      1      1      pass
U107      S204      A304      1465130523      1475130523      1485130523      A      ST405      1      1      0      1      1      pass
U120      S204      A304      1465230523      1485130523      1465230523      AU      ST401      1      0      1      1      1      pass
U104      S204      A304      1475130523      1465230523      1465130523      A      ST400      3      0      0      1      1      pass
U107      S204      A304      1494399262      1497026556      1497026556      AP      ST407      1      0      1      1      1      pass
U114      S205      A301      1494399262      1497026556      1525833562      E      ST414      3      1      0      1      1      pass
U119      S207      A303      1495130523      1465230523      1485130523      J      ST413      3      1      0      1      1      pass
U115      S207      A303      1465230523      1485130523      1475130523      AP      ST412      2      1      0      1      1      pass
U113      S207      A303      1465230523      1465230523      1475130523      E      ST409      1      0      0      1      1      pass
U117      S208      A304      1499630889      1525833562      1494399262      A      ST411      1      1      0      1      1      pass
U101      S208      A304      1475130523      1465230523      1475130523      A      ST411      1      0      0      1      1      pass
U103      S208      A304      1499630889      1494399262      1525833562      A      ST405      3      0      1      1      1      pass

```

e. Two folders “invalid” & “valid” created inside the processed\_dir in local.

f. Valid and invalid files created earlier were copied to the local directories created above

11) data\_analysis.sh -- Performs data analysis

- a. As per project requirements, 5 different analysis report to be created on the formatted, enhanced data.
- b. Five different hive tables were created and the analyzed formatted data loaded into them
- c. Top 10 stations – where maximum number of songs were played, which were liked by unique users.

```

hive> USE PROJECT;
OK
Time taken: 14.373 seconds
hive> CREATE TABLE IF NOT EXISTS top_10_stations
> (
>   station_id STRING,
>   total_distinct_songs_played INT,
>   distinct_user_count INT
> )
> PARTITIONED BY (batchid INT)
> ROW FORMAT DELIMITED
> FIELDS TERMINATED BY ','
> STORED AS TEXTFILE;
OK
Time taken: 0.585 seconds

```

```

hive> INSERT OVERWRITE TABLE top_10_stations
> PARTITION(batchid=1)
> SELECT
> station_id,
> COUNT(DISTINCT song_id) AS total_distinct_songs_played,
> COUNT(DISTINCT user_id) AS distinct_user_count
> FROM batched_data
> WHERE status='pass'
> AND batchid=1
> AND like1=1
> GROUP BY station_id
> ORDER BY total_distinct_songs_played DESC
> LIMIT 10;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X releases.
Query ID = acadgild_20190325224504_4a7bbac8-6282-41df-9306-a3c0c0e49571
Total jobs = 2
Launching Job 1 out of 2
Number of reduce tasks not specified. Estimated from input data size: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1553492063162_0015, Tracking URL = http://localhost:8088/proxy/application_1553492063162_0015/
Kill Command = /home/acadgild/install/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1553492063162_0015
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2019-03-25 22:45:40,413 Stage-1 map = 0%, reduce = 0%
2019-03-25 22:45:56,980 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.99 sec
2019-03-25 22:46:13,244 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 7.47 sec
MapReduce Total cumulative CPU time: 7 seconds 476 msec
Ended Job = job_1553492063162_0015
Launching Job 2 out of 2
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:

```

```
hive> select * from top_10_stations;
```

```
OK
```

|       |   |   |   |
|-------|---|---|---|
| ST412 | 2 | 2 | 1 |
| ST405 | 2 | 2 | 1 |
| ST414 | 1 | 1 | 1 |
| ST413 | 1 | 1 | 1 |
| ST411 | 1 | 1 | 1 |
| ST406 | 1 | 1 | 1 |
| ST401 | 1 | 1 | 1 |
| ST402 | 1 | 1 | 1 |

```
Time taken: 0.978 seconds, Fetched: 8 row(s)
```

- d. **UsersBehaviour** - Determine total duration of songs played by each type of user, where type of user can be 'subscribed' or 'unsubscribed'. An unsubscribed user is the one whose record is either not present in Subscribed\_users lookup table or has *subscription\_end\_date* earlier than the *timestamp* of the song played by him.

```

hive>
>
> CREATE TABLE IF NOT EXISTS users_behaviour
> (
>   user_type STRING,
>   duration INT
> )
> PARTITIONED BY (batchid INT)
> ROW FORMAT DELIMITED
> FIELDS TERMINATED BY ','
> STORED AS TEXTFILE;
OK
Time taken: 0.105 seconds

```

```

File Edit View Search Terminal Help
hive> INSERT OVERWRITE TABLE usersBehaviour
> PARTITION(batchid=1)
> SELECT
> CASE WHEN (su.user_id IS NULL OR CAST(ed.time_stamp AS DECIMAL(20,0)) > CAST(su.subscn_end_dt AS DECIMAL(20,0))) THEN
UNSUBSCRIBED'
> WHEN (su.user_id IS NOT NULL AND CAST(ed.time_stamp AS DECIMAL(20,0)) <= CAST(su.subscn_end_dt AS DECIMAL(20,0))) THEN
'SUBSCRIBED'
> ELSE AS user_type,
> SUM(DS(CAST(ed.end_ts AS DECIMAL(20,0))-CAST(ed.start_ts AS DECIMAL(20,0)))) AS duration
> FROM enriched_data ed
> LEFT OUTER JOIN subscribed_users su
> ON ed.user_id=su.user_id
> WHERE ed.status='pass'
> AND ed.batchid=1
> GROUP BY CASE WHEN (su.user_id IS NULL OR CAST(ed.time_stamp AS DECIMAL(20,0)) > CAST(su.subscn_end_dt AS DECIMAL(20,0)))
)) THEN 'UNSUBSCRIBED'
> WHEN (su.user_id IS NOT NULL AND CAST(ed.time_stamp AS DECIMAL(20,0)) <= CAST(su.subscn_end_dt AS DECIMAL(20,0))) THEN
'SUBSCRIBED' END;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execu
tive engine (i.e. spark, tez) or Hive 1.X releases.
Query ID : acadgild_2019032525034_413b4815-e7b0-4132-a230-22da73ef7b1d
Total jobs: 1
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/acadgild/install/hive/apache-hive-2.3.2-bin/lib/log4j-slf4j-impl-2.6.2.jar!org/slf4j
/impl/StaticLoggerBinder.class]
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/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
2019-03-25 22:50:53 Starting to launch local task to process map join; maximum memory = 518979584
2019-03-25 22:50:59 Dump the side-table for tag: 1 with group count: 15 into file: file:/tmp/acadgild/aa18e6c3-ba5f-45c6-
8306-cabed8e4886c/hive_2019-03-25_22-50-34_486_1192488398001868728-1/-local-10003/HashTable-Stage-2/MapJoin-mapfile01--.hasht
able
2019-03-25 22:50:59 Uploaded 1 File to file:/tmp/acadgild/aa18e6c3-ba5f-45c6-8306-cabed8e4886c/hive_2019-03-25_22-50-34_
486_1192488398001868728-1/-local-10003/HashTable-Stage-2/MapJoin-mapfile01--.hashtable (773 bytes)
2019-03-25 22:50:59 End of local task; Time Taken: 5.901 sec.
Execution completed successfully
MapredLocal task succeeded
Launching Job 1 out of 1
Launching Job 1 out of 1
Number of reduce tasks not specified. Estimated from input data size: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1553492063162_0017, Tracking URL = http://localhost:8088/proxy/application_1553492063162_0017/
Kill Command = /home/acadgild/install/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1553492063162_0017
Hadoop job information for Stage-2: number of mappers: 1; number of reducers: 1
2019-03-25 22:51:20,961 Stage-2 map = %, reduce = 0%
2019-03-25 22:51:36,181 Stage-2 map = 100%, reduce = 0%, Cumulative CPU 4.73 sec
2019-03-25 22:51:53,704 Stage-2 map = 100%, reduce = 100%, Cumulative CPU 10.0 sec
MapReduce Total cumulative CPU time: 10 seconds 0 msec
Ended Job = job_1553492063162_0017
Loading data to table project.usersBehaviour partition (batchid=1)
MapReduce Jobs Launched:
Stage-Stage-2: Map: 1 Reduce: 1 Cumulative CPU: 10.0 sec HDFS Read: 26733 HDFS Write: 132 SUCCESS
Total MapReduce CPU Time Spent: 10 seconds 0 msec
OK
Time taken: 82.949 seconds
hive> select * from usersBehaviour;
OK
SUBSCRIBED      29800000      1
UNSUBSCRIBED    255655545      1
Time taken: 0.58 seconds, Fetched: 2 row(s)

```

e. Top 10 connected artists - Connected artists are those whose songs are most listened by the unique users who follow them.

```

hive> CREATE TABLE IF NOT EXISTS connected_artists
> (
>   artist_id STRING,
>   user_count INT
> )
> PARTITIONED BY (batchid INT)
> ROW FORMAT DELIMITED
> FIELDS TERMINATED BY ','
> STORED AS TEXTFILE;
OK
Time taken: 0.082 seconds

```

```

Hadoop 2.6.1 [Running] - Oracle VM VirtualBox
Applications Places System acadgild@localhost:~ 
File Edit View Search Terminal Help
Time taken: 0.082 seconds
hive> INSERT OVERWRITE TABLE connected_artists
> PARTITION(batchid=1)
> SELECT
> ua.artist_id,
> COUNT(DISTINCT ua.user_id) AS user_count
> FROM
> (
>   SELECT user_id, artist_id FROM users_artists
>   LATERAL VIEW explode(artists_array) artists AS artist_id
> ) ua
> INNER JOIN
> (
>   SELECT artist_id, song_id, user_id
>   FROM enriched_data
>   WHERE status='pass'
>   AND batchid=1
> ) ed
> ON ua.artist_id=ed.artist_id
> AND ua.user_id=ed.user_id
> GROUP BY ua.artist_id
> ORDER BY user_count DESC
> LIMIT 10;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X releases.
Query ID = acadgild_2019032525450_0ce63567-c0dd-4033-bac8-68b8a3f581e8
Total jobs = 2
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/acadgild/install/hive/apache-hive-2.3.2-bin/lib/log4j-slf4j-impl-2.6.2.jar!/org/slf4j/impl/StaticLoggerBinder.class]
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/StaticLoggerBinder.class]
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory].
2019-03-25 22:55:10  Starting to launch local_task to process map joins... maximum memory = 518979584
2019-03-25 22:55:13      Dump the side-table for tag: 0 with group count: 33 into file: file:/tmp/acadgild/a18e6c3-ba5f-45c6-8306-cabed8e4886c/hive_2019-03-25-22-54-50_412_2955797518108744269-1-local-10004/HashTable-Stage-2/MapJoin-mapfile10--.hashtable
2019-03-25 22:55:13      Uploaded 1 File to: file:/tmp/acadgild/a18e6c3-ba5f-45c6-8306-cabed8e4886c/hive_2019-03-25 22-54-50_412_2955797518108744269-1-local-10004/HashTable-Stage-2/MapJoin-mapfile10--.hashtable

```

```

Hadoop 2.6.1 [Running] - Oracle VM VirtualBox
Applications Places System acadgild@localhost:~ 
File Edit View Search Terminal Help
set mapreduce.job.reduces=<numbers>
Starting Job: job_1553492063162_0018, Tracking URL = http://localhost:8088/proxy/application_1553492063162_0018/
Kill Command = /home/acadgild/install/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1553492063162_0018
Hadoop job information for Stage-2: number of mappers: 1; number of reducers: 1
2019-03-25 22:55:32,898 Stage-2 map = 0%, reduce = 0%
2019-03-25 22:55:48,438 Stage-2 map = 100%, reduce = 0%, Cumulative CPU 4.43 sec
2019-03-25 22:56:04,879 Stage-2 map = 100%, reduce = 100%, Cumulative CPU 7.87 sec
MapReduce Total cumulative CPU time: 7 seconds 870 msec
Ended Job = job_1553492063162_0018
Launching Job 2 out of 2
Number of reduce tasks determined at compile time: 1
In order to change the average bytes for a reducer (in bytes):
  set hive.mapred.reducer.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<numbers>
Starting Job = job_1553492063162_0019, Tracking URL = http://localhost:8088/proxy/application_1553492063162_0019/
Kill Command = /home/acadgild/install/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1553492063162_0019
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 1
2019-03-25 22:56:25,106 Stage-3 map = 0%, reduce = 0%
2019-03-25 22:56:37,392 Stage-3 map = 100%, reduce = 0%, Cumulative CPU 1.94 sec
2019-03-25 22:56:52,544 Stage-3 map = 100%, reduce = 100%, Cumulative CPU 6.4 sec
MapReduce Total cumulative CPU time: 6 seconds 400 msec
Ended Job = job_1553492063162_0019
Loading data to table project.connected_artists partition (batchid=1)
MapReduce Jobs Launched:
Stage-Stage-2: Map: 1 Reduce: 1   Cumulative CPU: 7.87 sec   HDFS Read: 15675 HDFS Write: 165 SUCCESS
Stage-Stage-3: Map: 1 Reduce: 1   Cumulative CPU: 6.4 sec    HDFS Read: 6632 HDFS Write: 112 SUCCESS
Total MapReduce CPU Time Spent: 14 seconds 270 msec
OK
Time taken: 126.51 seconds
hive> select * from connected_artists;
OK
A301    2      1
A384    1      1
A302    1      1
Time taken: 0.526 seconds, Fetched: 3 row(s)
hive> 
```

f. Top 10 royalty songs - who have generated the maximum revenue. Royalty applies to a song only if it was *liked* or was *completed successfully* or both.

```

hive> CREATE TABLE IF NOT EXISTS top_10_royalty_songs
> (
>   song_id STRING,
>   duration INT
> )
> PARTITIONED BY (batchid INT)
> ROW FORMAT DELIMITED
> FIELDS TERMINATED BY ','
> STORED AS TEXTFILE;
OK
Time taken: 0.046 seconds

```

```

File Edit View Search Terminal Help
hive> INSERT OVERWRITE TABLE top_10_royalty_songs
> PARTITION(batchid=1)
> SELECT song_id,
> SUM(ABS(CAST(end_ts AS DECIMAL(20,0))-CAST(start_ts AS DECIMAL(20,0)))) AS duration
> FROM enriched_data
> WHERE status='pass'
> AND batchid=1
> AND (like1=1 OR song_end_type=0)
> GROUP BY song_id
> ORDER BY duration DESC
> LIMIT 10;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X releases.
Query ID = acadgild_20190325230024_92469e6a-37f3-41d4-b85a-ae010eb05295
Total jobs = 2
Launching Job 1 out of 2
Number of reduce tasks not specified. Estimated from input data size: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1553492063162_0020, Tracking URL = http://localhost:8088/proxy/application_1553492063162_0020/
Kill Command = /home/acadgild/hadoop/install/hadoop-2.6.5/bin/hadoop job -kill job_1553492063162_0020
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2019-03-25 23:00:42,873 Stage-1 map = 0%, reduce = 0%
2019-03-25 23:00:57,013 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 4.68 sec
2019-03-25 23:01:14,965 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 8.08 sec
MapReduce Total cumulative CPU time: 8 seconds 80 msec
Ended Job = job_1553492063162_0020
Launching Job 2 out of 2
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
hive>
    select * from top_10_royalty_songs
  ;
OK
S208      34038633      1
S207      29900000      1
S205      28807006      1
S203      26202673      1
S204      10000000      1
S201      99000000      1
S202      7858921       1
Time taken: 0.334 seconds, Fetched: 7 row(s)
hive> ■

```

g. Top 10 unsubscribed users - top 10 unsubscribed users who listened to the songs for the longest duration.

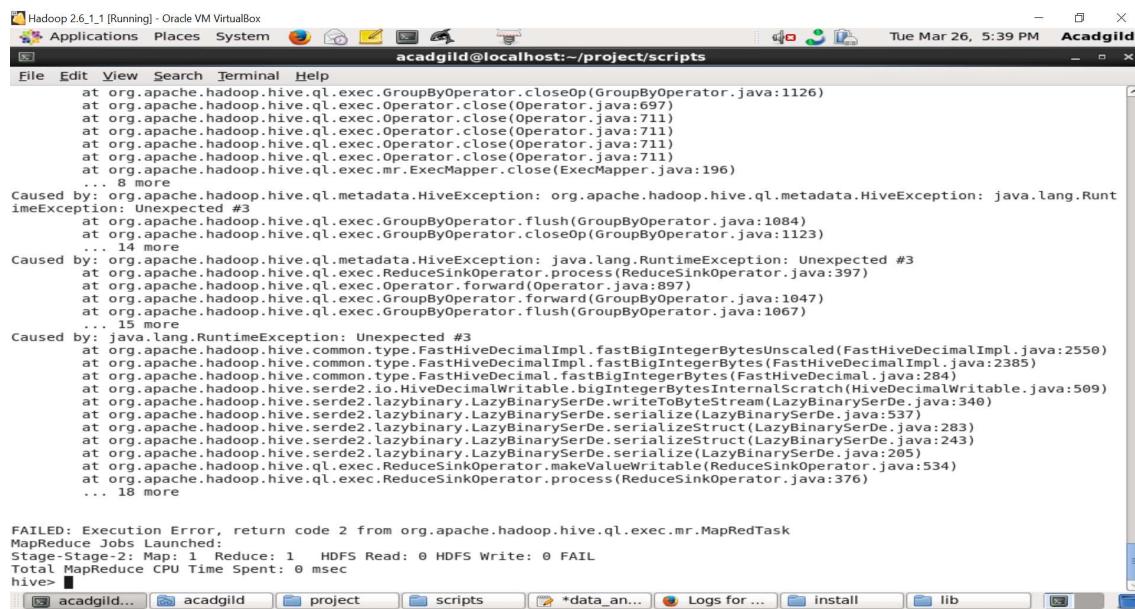
```

hive> CREATE TABLE IF NOT EXISTS top_10_unsubscribed_users
  > (
  >   user_id STRING,
  >   duration INT
  > )
  > PARTITIONED BY (batchid INT)
  > ROW FORMAT DELIMITED
  > FIELDS TERMINATED BY ','
  > STORED AS TEXTFILE;
OK
Time taken: 0.099 seconds

```

- h. The insert statement given below could not be executed due to Unexpected error #3 in Java

```
INSERT OVERWRITE TABLE top_10_unsubscribed_users
PARTITION(batchid=${hiveconf:batchid})
SELECT
ed.user_id,
SUM(ABS(CAST(ed.end_ts AS DECIMAL(20,0))-CAST(ed.start_ts AS DECIMAL(20,0)))) AS duration
FROM enriched_data ed
LEFT OUTER JOIN subscribed_users su
ON ed.user_id=su.user_id
WHERE ed.status='pass'
AND ed.batchid=${hiveconf:batchid}
AND (su.user_id IS NULL OR (CAST(ed.timestamp AS DECIMAL(20,0)) > CAST(su.subscn_end_dt AS
DECIMAL(20,0))))
GROUP BY ed.user_id
ORDER BY duration DESC
LIMIT 10;
```



The screenshot shows a terminal window titled "Hadoop 2.6.1\_1 [Running] - Oracle VM VirtualBox". The command entered was:

```
hive> INSERT OVERWRITE TABLE top_10_unsubscribed_users
PARTITION(batchid=${hiveconf:batchid})
SELECT
ed.user_id,
SUM(ABS(CAST(ed.end_ts AS DECIMAL(20,0))-CAST(ed.start_ts AS DECIMAL(20,0)))) AS duration
FROM enriched_data ed
LEFT OUTER JOIN subscribed_users su
ON ed.user_id=su.user_id
WHERE ed.status='pass'
AND ed.batchid=${hiveconf:batchid}
AND (su.user_id IS NULL OR (CAST(ed.timestamp AS DECIMAL(20,0)) > CAST(su.subscn_end_dt AS
DECIMAL(20,0))))
GROUP BY ed.user_id
ORDER BY duration DESC
LIMIT 10;
```

The output shows an error message:

```
FAILED: Execution Error, return code 2 from org.apache.hadoop.hive.ql.exec.mr.MapRedTask
MapReduce Jobs Launched:
Stage-Stage-2: Map: 1 Reduce: 1 HDFS Read: 0 HDFS Write: 0 FAIL
Total MapReduce CPU Time Spent: 0 msec
hive> 
```

Below the terminal window, there is a horizontal toolbar with several icons and labels: acadgild... acadgild project scripts \*data\_an... Logs for ... install lib.

- i. Created the objects in mysql as per the schema provided

```

Hadoop 2.6.1_1 [Running] - Oracle VM VirtualBox
Applications Places System acadgild@localhost:~ 
File Edit View Search Terminal Help
mysql> CREATE DATABASE IF NOT EXISTS project;
Query OK, 1 row affected (0.00 sec)

mysql> USE project;
Database changed
mysql> CREATE TABLE IF NOT EXISTS top_10_stations
-> (
->   station_id VARCHAR(50),
->   total_distinct_songs_played INT,
->   distinct_user_count INT
-> );
Query OK, 0 rows affected (0.02 sec)

mysql> CREATE TABLE IF NOT EXISTS users_behaviour
-> (
->   user_type VARCHAR(50),
->   duration BIGINT
-> );
Query OK, 0 rows affected (0.02 sec)

mysql>
mysql> CREATE TABLE IF NOT EXISTS connected_artists
-> (
->   artist_id VARCHAR(50),
->   user_count INT
-> );
Query OK, 0 rows affected (0.02 sec)

mysql>
mysql> CREATE TABLE IF NOT EXISTS top_10_royalty_songs
-> (
->   song_id VARCHAR(50),
->   duration BIGINT
-> );
Query OK, 0 rows affected (0.02 sec)

mysql>
mysql> CREATE TABLE IF NOT EXISTS top_10_unsubscribed_users
-> (
->   user_id VARCHAR(50),
->   duration BIGINT
-> );
Query OK, 0 rows affected (0.02 sec)

mysql>
mysql> commit;
Query OK, 0 rows affected (0.00 sec)

mysql>
Display all 763 possibilities? (y or n)
mysql> 

```

j. Using the sqoop scripts provided, data was loaded into mysql

Export 1

## Export 2

```

Hadoop 2.6.1_1 [Running] - Oracle VM VirtualBox
Applications Places System acadgild@localhost:~ File Edit View Search Terminal Help
[acadgild@localhost ~]$ sqoop export --connect jdbc:mysql://localhost/project --username 'root' --password 'Vramasub#1' --table users behaviour --export-dir hdfs://localhost:8020/user/hive/warehouse/project.db/usersBehaviour/batchid=1 --input-fields terminated-by ',' -m 1
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 SHCAT_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.
19/03/27 19:30:48 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6
19/03/27 19:30:48 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
19/03/27 19:30:49 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
19/03/27 19:30:49 INFO tool.CodeGenTool: Beginning code generation
Wed Mar 27 19:30:49 IST 2019 WARN: Establishing SSL connection without server's identity verification is not recommended. According to MySQL 5.5.45+, 5.6.26+ and 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 verifyServerCertificate property is set to 'false'. You need either to explicitly disable SSL by setting useSSL=false, or set useSSL=true and provide truststore for server certificate verification.
19/03/27 19:30:51 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `users behaviour` AS t LIMIT 1
19/03/27 19:30:51 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `users behaviour` AS t LIMIT 1
19/03/27 19:30:51 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /home/acadgild/install/hadoop/hadoop-2.6.5
Note: /tmp/sqoop-acadgild/compile/97d4bb021273bd54d4e636aab15f21018/usersBehaviour.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
19/03/27 19:30:54 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-acadgild/compile/97d4bb021273bd54d4e636aab15f21018/usersBehaviour.jar
19/03/27 19:30:54 INFO mapreduce.ExportJobBase: Beginning export of usersBehaviour
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/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/acadgild/install/hbase/hbase-1.2.6/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
19/03/27 19:30:55 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
19/03/27 19:30:55 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar
19/03/27 19:30:56 INFO Configuration.deprecation: mapred.reduce.tasks.speculative.execution is deprecated. Instead, use mapreduce.reduce.speculative
19/03/27 19:30:57 INFO Configuration.deprecation: mapred.map.tasks.speculative.execution is deprecated. Instead, use mapreduce.map.tasks.speculative.execution
[acadgild@localhost ~]$ How to... acadgild... Trash Compu... acadgild project scripts mysql1... acadgild...
Hadoop 2.6.1_1 [Running] - Oracle VM VirtualBox
Applications Places System acadgild@localhost:~ File Edit View Search Terminal Help
[acadgild@localhost ~]$ select * from usersBehaviour;
+-----+-----+-----+
| station_id | total_distinct_songs_played | distinct_user_count |
+-----+-----+-----+
| ST412      | 2                   | 2                   |
| ST405      | 2                   | 2                   |
| ST414      | 1                   | 1                   |
| ST413      | 1                   | 1                   |
| ST411      | 1                   | 1                   |
| ST406      | 1                   | 1                   |
| ST401      | 1                   | 1                   |
| ST402      | 1                   | 1                   |
+-----+-----+-----+
8 rows in set (0.00 sec)

mysql> select * from usersBehaviour;
+-----+-----+
| user_type | duration |
+-----+-----+
| SUBSCRIBED | 29800000 |
| UNSUBSCRIBED | 255655545 |
+-----+-----+
2 rows in set (0.00 sec)

mysql> [acadgild@localhost ~]$ How to... acadgild... Trash Compu... acadgild project scripts mysql1... acadgild...

```

### Export 3

```
[acadgild@localhost ~]$ sqoop export --connect jdbc:mysql://localhost/project --username 'root' --password 'Vramasub#1' --table connected_artists --export-dir hdfs://localhost:8020/user/hive/warehouse/project.db/connected_artists/batchid=1 --input-fields-terminated-by ',' -m 1
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 SHCAT_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.
19/03/27 19:33:08 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6
19/03/27 19:33:08 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
19/03/27 19:33:09 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
19/03/27 19:33:09 INFO tool.CodeGenTool: Beginning code generation
Wed Mar 27 19:33:09 IST 2019 WARN: Establishing SSL connection without server's identity verification is not recommended. According to MySQL 5.5.45+, 5.6.26+ and 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 verifyServerCertificate property is set to 'false'. You need either to explicitly disable SSL by setting useSSL=false, or set useSSL=true and provide truststore for server certificate verification.
19/03/27 19:33:11 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `connected_artists` AS t LIMIT 1
19/03/27 19:33:11 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `connected_artists` AS t LIMIT 1
19/03/27 19:33:11 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /home/acadgild/install/hadoop/hadoop-2.6.5
Note: /tmp/sqoop-acadgild/compile/49c077259870e9c3abaa26165b3115d8/connected_artists.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
19/03/27 19:33:15 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-acadgild/compile/49c077259870e9c3abaa26165b3115d8/connected_artists.jar
19/03/27 19:33:15 INFO mapreduce.ExportJobBase: Beginning export of connected_artists
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/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/acadgild/install/hbase/hbase-1.2.6/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
19/03/27 19:33:15 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
19/03/27 19:33:15 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar
19/03/27 19:33:17 INFO Configuration.deprecation: mapred.reduce.tasks.speculative.execution is deprecated. Instead, use mapreduce.reduce.speculative
[acadgild@localhost ~]$ How to... acadgild... Trash Compu... acadgild project scripts mysql1... acadgild...
[acadgild@localhost ~]$ Hadoop 2.6.1_1 [Running] - Oracle VM VirtualBox
[acadgild@localhost ~]$ Applications Places System
[acadgild@localhost ~]$ File Edit View Search Terminal Help
[acadgild@localhost ~]$ mysql> select * from top_10_stations;
Empty set (0.00 sec)

mysql> select * from top_10_stations;
+-----+-----+-----+
| station_id | total_distinct_songs_played | distinct_user_count |
+-----+-----+-----+
| ST412      | 2                | 2                  |
| ST405      | 2                | 2                  |
| ST414      | 1                | 1                  |
| ST413      | 1                | 1                  |
| ST411      | 1                | 1                  |
| ST406      | 1                | 1                  |
| ST401      | 1                | 1                  |
| ST402      | 1                | 1                  |
+-----+-----+-----+
8 rows in set (0.00 sec)

mysql> select * from users_beaviour;
+-----+-----+
| user_type | duration |
+-----+-----+
| SUBSCRIBED | 29800000 |
| UNSUBSCRIBED | 255655545 |
+-----+-----+
2 rows in set (0.00 sec)

mysql> mysql> select * from connected_artists;
+-----+-----+
| artist_id | user_count |
+-----+-----+
| A301      | 2          |
| A304      | 1          |
| A302      | 1          |
+-----+-----+
3 rows in set (0.00 sec)

mysql> [acadgild@localhost ~]$ How to... acadgild... Trash Compu... acadgild project scripts mysql1... acadgild...

```

## Export 4

```

Hadoop 2.6.1_1 [Running] - Oracle VM VirtualBox
Applications Places System acadgild@localhost:~ File Edit View Search Terminal Help
[acadgild@localhost ~]$ sqoop export --connect jdbc:mysql://localhost/project --username 'root' --password 'Vramasub#1' --tab le top_10_royalty_songs --export-dir hdfs://localhost:8020/user/hive/warehouse/project.db/top_10_royalty_songs/batchid=1 --in put-fields-terminated-by ',' -m 1
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 SHCAT_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 f ail.
Please set SACCUMULO_HOME to the root of your Accumulo installation.
19/03/27 19:34:48 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6
19/03/27 19:34:48 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
19/03/27 19:34:48 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
19/03/27 19:34:48 INFO tool.CodeGenTool: Beginning code generation
Wed Mar 27 19:34:48 IST 2019 WARN: Establishing SSL connection without this server's identity verification is not recommended. According to MySQL 5.5.45+, 5.6.26+ and 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 verifyServerCertificate property is set to 'false'. You need either to explicitly disable SSL by setting useSSL=false, or set useSSL=true and provide truststore for server certificate verification.
19/03/27 19:34:50 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `top_10_royalty_songs` AS t LIMIT 1
19/03/27 19:34:50 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `top_10_royalty_songs` AS t LIMIT 1
19/03/27 19:34:50 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /home/acadgild/install/hadoop/hadoop-2.6.5
Note: /tmp/sqoop-acadgild/compile/4a5415cb216eefb038cd951986844bd7/top_10_royalty_songs.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
19/03/27 19:34:53 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-acadgild/compile/4a5415cb216eefb038cd951986844bd7/top_10_royalty_songs.jar
19/03/27 19:34:53 INFO mapreduce.ExportJobBase: Beginning export of top_10_royalty_songs
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/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/acadgild/install/hbase/hbase-1.2.6/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
19/03/27 19:34:54 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
19/03/27 19:34:54 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar
19/03/27 19:34:54 INFO Configuration.deprecation: mapred.reduce.tasks.speculative.execution is deprecated. Instead, use mapreduce.reduce.speculative
[ How to... ] [ acadgil... ] [ Trash ] [ Compu... ] [ acadgild ] [ project ] [ scripts ] [ mysql1... ] [ acadgil... ] [ Applications ] [ Places ] [ System ] [ Wed Mar 27, 7:35 PM ] [ Acadgild ]
Hadoop 2.6.1_1 [Running] - Oracle VM VirtualBox
Applications Places System acadgild@localhost:~ File Edit View Search Terminal Help
[acadgild@localhost ~]$ ST401 | 1 | 1 |
| ST402 | 1 | 1 |
+-----+-----+-----+
8 rows in set (0.00 sec)

mysql> select * from users_behaviour;
+-----+-----+
| user_type | duration |
+-----+-----+
| SUBSCRIBED | 29800000 |
| UNSUBSCRIBED | 255655545 |
+-----+-----+
2 rows in set (0.00 sec)

mysql> mysql> select * from connected_artists;
+-----+-----+
| artist_id | user_count |
+-----+-----+
| A301 | 2 |
| A304 | 1 |
| A302 | 1 |
+-----+-----+
3 rows in set (0.00 sec)

mysql> select * from top_10_royalty_songs;
+-----+-----+
| song_id | duration |
+-----+-----+
| S208 | 34938633 |
| S207 | 29800000 |
| S205 | 28800000 |
| S203 | 26202673 |
| S204 | 10000000 |
| S201 | 99000000 |
| S202 | 7858921 |
+-----+-----+
7 rows in set (0.00 sec)

mysql> [ How to... ] [ acadgil... ] [ Trash ] [ Compu... ] [ acadgild ] [ project ] [ scripts ] [ mysql1... ] [ acadgil... ] [ Applications ] [ Places ] [ System ] [ Wed Mar 27, 7:36 PM ] [ Acadgild ]

```

## Export 5:

```
Hadoop 2.6.1_1 [Running] - Oracle VM VirtualBox
Applications Places System acadgild@localhost:~ 
File Edit View Search Terminal Help
[acadgild@localhost ~]$ sqoop export --connect jdbc:mysql://localhost/project --username 'root' --password 'Vramasub#1' --table top_10_unsubscribed_users --export-dir hdfs://localhost:8020/user/hive/warehouse/project.db/top_10_unsubscribed_users/batch-id=1 --input-fields-terminated-by ',' -m 1
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.
19/03/27 19:41:22 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6
19/03/27 19:41:22 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
19/03/27 19:41:23 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
19/03/27 19:41:23 INFO tool.CodeGenTool: Beginning code generation
Wed Mar 27 19:41:23 IST 2019 WARN: Establishing SSL connection without server's identity verification is not recommended. According to MySQL 5.5.45+, 5.6.26+ and 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 verifyServerCertificate property is set to 'false'. You need either to explicitly disable SSL by setting useSSL=false, or set useSSL=true and provide truststore for server certificate verification.
19/03/27 19:41:24 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `top_10_unsubscribed_users` AS t LIMIT 1
19/03/27 19:41:24 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `top_10_unsubscribed_users` AS t LIMIT 1
19/03/27 19:41:25 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /home/acadgild/install/hadoop/hadoop-2.6.5
Note: /tmp/sqoop-acadgild/compile/970ef20f864ccedd47facae612d40fb/top_10_unsubscribed_users.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
19/03/27 19:41:28 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-acadgild/compile/970ef20f864ccedd47facae612d40fb/top_10_unsubscribed_users.jar
19/03/27 19:41:28 INFO mapreduce.ExportJobBase: Beginning export of top_10_unsubscribed_users
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/jul/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/acadgild/install/hbase/hbase-1.2.6/lib/slf4j-log4j12-1.7.5.jar!]/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
19/03/27 19:41:28 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
19/03/27 19:41:28 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar
19/03/27 19:41:30 WARN mapreduce.ExportJobBase: Input path hdfs://localhost:8020/user/hive/warehouse/project.db/top_10_unsubscribed_users/batchId=1 contains no files
```

Crontab was missing in my VM. Using the following command crontab was installed in my vm

`sudo yum install cronie`

Schedule Crontab to run the wrapper.sh every three hours

A screenshot of a Linux desktop environment. At the top, there's a menu bar with "Applications", "Places", "System", and "Help". The title bar of the active window says "acadgild@localhost:~/install/hive/apache-hive-2.3.2-bin/bin". The window contains a terminal session with the following text:

```
* */3 * * * /home/acadgild/project/scripts/wrapper.sh
```

The desktop background is light blue with a subtle grid pattern. The taskbar at the bottom shows icons for various applications like a browser, file manager, and system tools.

Once the wrapper started executing, new log file was created with batch\_id 3

Hadoop 2.6.1\_1 [Running] - Oracle VM VirtualBox

Applications Places System

File Edit View Search Tools Documents Help

Open Save Undo Redo Cut Copy Paste Find Replace Select All

Solution\_execution.txt wrapper.sh \*data\_enrichment\_filtering\_schema.sh log\_batch\_3

Starting daemons  
Creating LookUp Tables

Plain Text Tab Width: 8 Ln 1, Col 1 INS

sh... Tr... Co... pr... sc... So... ... ec... So... sc... ac... lo... logs

Hadoop 2.6.1\_1 [Running] - Oracle VM VirtualBox

Applications Places System

File Edit View Terminal Help

Wed Mar 27, 8:43 PM Acadgild

acadgild@localhost:~/project/scripts

```
2019-03-27 20:43:03,103 WARN [main] util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/acadgild/install/hbase/hbase-1.2.6/lib.slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]
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/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
HBase Shell; enter 'help<RETURN>' for list of supported commands.
Type "exit<RETURN>" to leave the HBase Shell
Version 1.2.6, rUnknown, Mon May 29 02:25:32 CDT 2017

put 'subscribed-users', 'U113', 'subscn:startdt', '1465230523'
0 row(s) in 1.3240 seconds

2019-03-27 20:43:23,571 WARN [main] util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/acadgild/install/hbase/hbase-1.2.6/lib.slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]
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/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
HBase Shell; enter 'help<RETURN>' for list of supported commands.
Type "exit<RETURN>" to leave the HBase Shell
Version 1.2.6, rUnknown, Mon May 29 02:25:32 CDT 2017

put 'subscribed-users', 'U113', 'subscn:enddt', '1485130523'
0 row(s) in 1.2630 seconds

2019-03-27 20:43:44,219 WARN [main] util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
```

in... Tr... Co... pr... sc... So... ... ec... So... sc... ac... wr... logs

Hadoop 2.6.1\_1 [Running] - Oracle VM VirtualBox

Applications Places System

Wed Mar 27, 8:38 PM Acadgild

File Edit View Search Tools Documents Help

Open Save Undo Redo Cut Copy Paste Find Replace

Solution\_execution.txt wrapper.sh \*data\_enrichment\_filtering\_schema.sh log\_batch\_3

Starting daemons  
Creating LookUp Tables  
Populating LookUp Tables

Plain Text Tab Width: 8 Ln 2, Col 23

INS

sh... Tr... Co... pr... sc... So... ... ec... So... sc... ac... lo... logs