**Task1**

Dataset is sample data of songs heard by users on an online streaming platform. The Description of data set attached in musicdata.txt is as follows: -

1st Column - UserId

2nd Column - TrackId

3rd Column - Songs Share status (1 for shared, 0 for not shared)

4th Column - Listening Platform (Radio or Web - 0 for radio, 1 for web)

5th Column - Song Listening Status (0 for skipped, 1 for fully heard)

**Task 1.1**

Write Map Reduce program for finding the number of unique listeners in the data set.



hadoop jar UniqueListeners.jar /Pradeep/Assignment3/musicdata.txt /Pradeep/Assignment3/UniqueListerners\_Output

hadoop 🡪To specify execute the command in HDFS

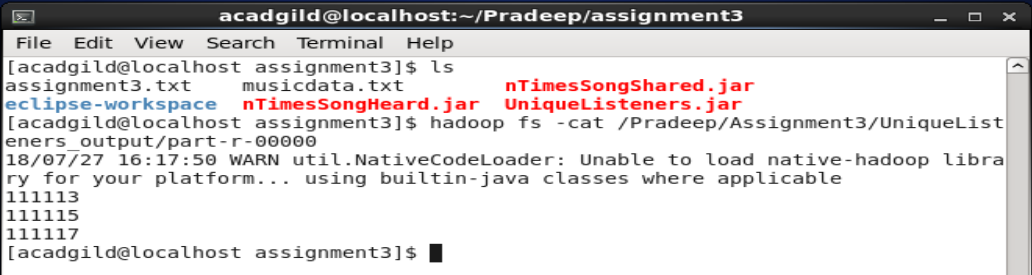
jar 🡪To specify we are going to execute the jar file

UniqueListeners.jar 🡪 Map Reduce program which will list the unique listeners in the given input file

/Pradeep/Assignment3/musicdata.txt 🡪 Path of the input file to be processed in HDFS

/Pradeep/Assignment3/UniqueListerners\_Output 🡪 Path of the output where the output of the result file to be stored

**OUTPUT:**



**Task 1.2**

Write a Map Reduce program to find the number of times a song was heard fully.



hadoop jar UniqueListeners.jar /Pradeep/Assignment3/musicdata.txt /Pradeep/Assignment3/UniqueListerners\_Output

hadoop 🡪To specify execute the command in HDFS

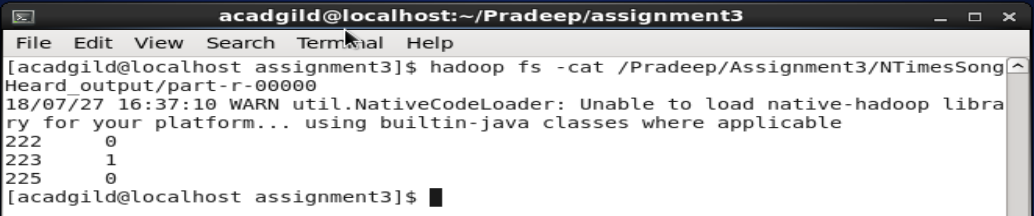
jar 🡪To specify we are going to execute the jar file

UniqueListeners.jar 🡪 Map Reduce program which will list the unique listeners in the given input file

/Pradeep/Assignment3/musicdata.txt 🡪 Path of the input file to be processed in HDFS

/Pradeep/Assignment3/UniqueListerners\_Output 🡪 Path of the output where the output of the result file to be stored

**OUTPUT:**



**Task 1.3**

Write a Map Reduce program to find the number of times a song was shared.



hadoop jar UniqueListeners.jar /Pradeep/Assignment3/musicdata.txt /Pradeep/Assignment3/UniqueListerners\_Output

hadoop 🡪To specify execute the command in HDFS

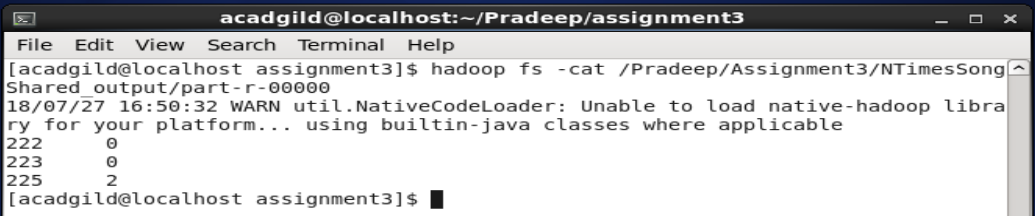
jar 🡪To specify we are going to execute the jar file

UniqueListeners.jar 🡪 Map Reduce program which will list the unique listeners in the given input file

/Pradeep/Assignment3/musicdata.txt 🡪 Path of the input file to be processed in HDFS

/Pradeep/Assignment3/UniqueListerners\_Output 🡪 Path of the output where the output of the result file to be stored

**OUTPUT:**



**Task 2**

**Task 2.1**

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

I have created the table **Person** in MySql assignment3db with the following command:

Create table Person(Person\_id int not null, Lname varchar(50), Fname varchar(50),

Area varchar(50), City varchar(50), PRIMARY KEY (person\_id) );

Then I have inserted the four rows into this table with the following commands:

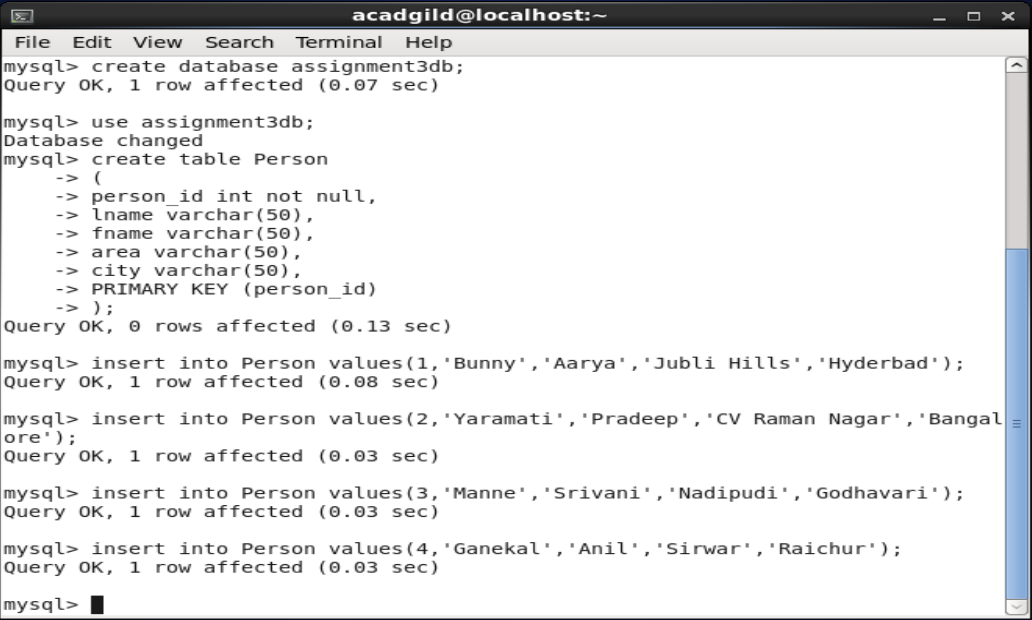
Insert into Person values(1,’Bunny’,’Aarya’,’Jubli Hills’,’Hyderabad’);

Insert into Person values(2,’Yaramati’,’Pradeep’,’CV Raman Nagar’,’Bangalore’);

Insert into Person values(3,’Manne’,’Srivani’,’Nadipudi’,’Godhavari’);

Insert into Person values(4,’Ganekal’,’Anil’,’Sirwar’,’Raichur’);

Commit;

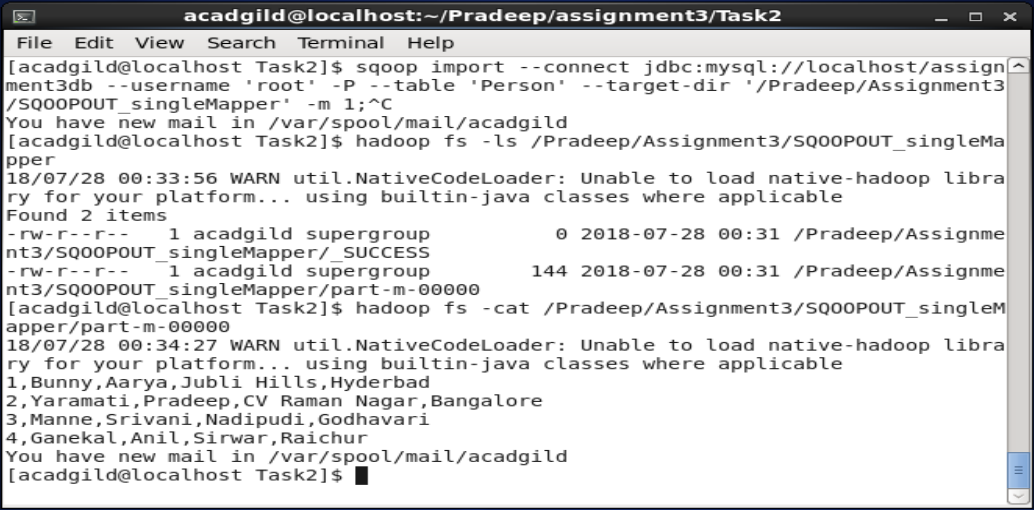


Now I have executed the following sqoop import command to fetch this data into HDFS path /Pradeep/Assignment3/SQOOPOUT\_SingleMapper folder:

Sqoop import –connect jdbc:mysql://localhost/assignment3db –username root –P –table ‘Person’ –target-dir ‘/Pradeep/Assignment3/SQOOPOUT\_singleMapper’ –m 1;

Hadoop fs –ls /Pradeep/Assignment3/SQOOPOUT\_singleMapper/

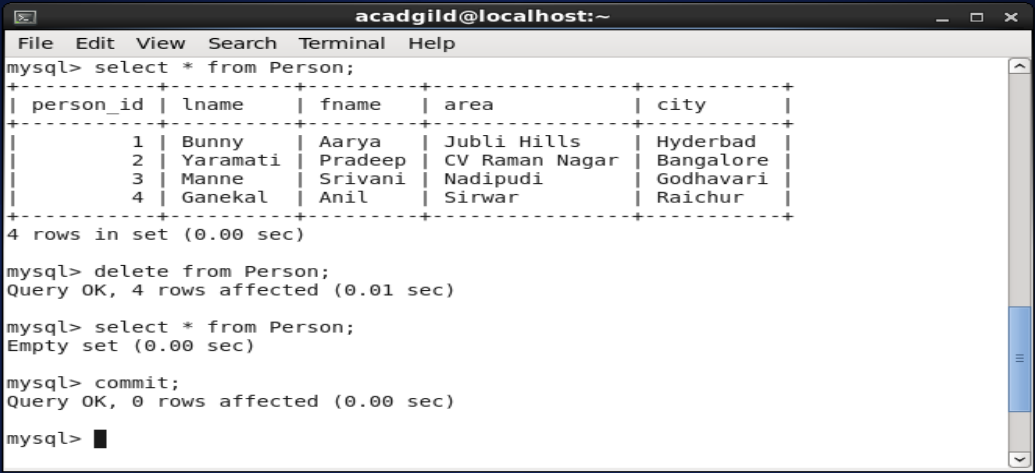
Hadoop fs –cat /Pradeep/Assignment3/SQOOPOUT\_singleMapper/part-m-00000



Now, I have went back to the MySql prompt and deleted the records present in the Person table with the following command:

Delete from Person;

Commit;

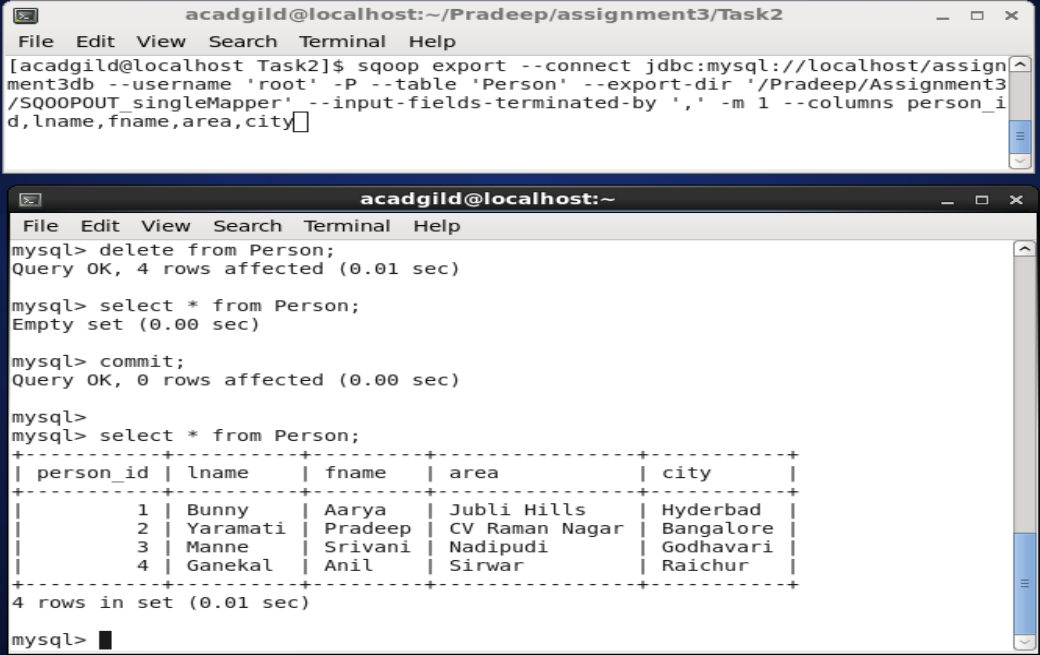


Now, I am going to use sqoop export command to transfer the data that we have gathered during the sqoop import command in /Pradeep/Assignment3/SQOOPOUT\_singleMapper to MySql database assignment3db table Person as the following command:

Sqoop export –connect jdbc:mysql://localhost/assignment3db –username ‘root’ –P –table ‘Person’ –export-dir ‘/Pradeep/Assignment3/SQOOPOUT\_singleMapper’ –input-fields-terminated-by ‘,’ –m 1 –columns person\_id,lname,fname,area,city

When I execute the following command I am able to see the contents of HDFS file are copied to the MySql database table.

Select \* from Person;



**Task 2.2**

Use Sqoop tool to import data present in SQOOPOUT folder made while demo of Import table with parameter person\_id =3.

I have executed the following command to fetch a particular record from the data base into the HDFS file system:

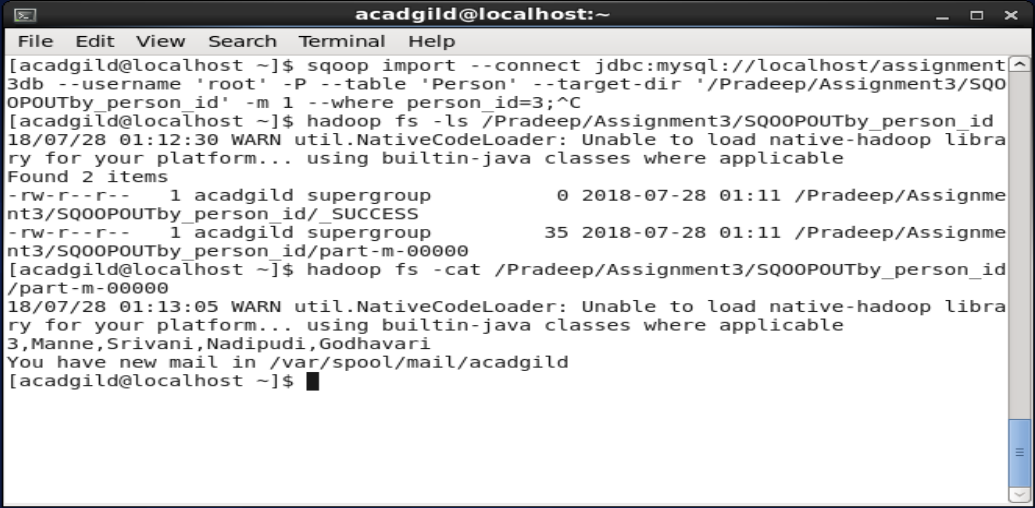
Sqoop import –connect jdbc:mysql://localhost/assignment3db –table Person –username ‘root’ –P –target-dir ‘/Pradeep/Assignment3/SQOOPOUTby\_person\_id’ –m 1 –where person\_id=3

Then I have checked the content of the result folder with the following command:

Hadoop fs –ls /Pradeep/Assignment3/SQOOPOUTby\_person\_id

Hadoop fs –cat /Pradeep/Assignment3/SQOOPOUTby\_person\_id/part-m-00000

The output file contains the desired record whose person\_id=3 as shown below:



Thank you,

Pradeep