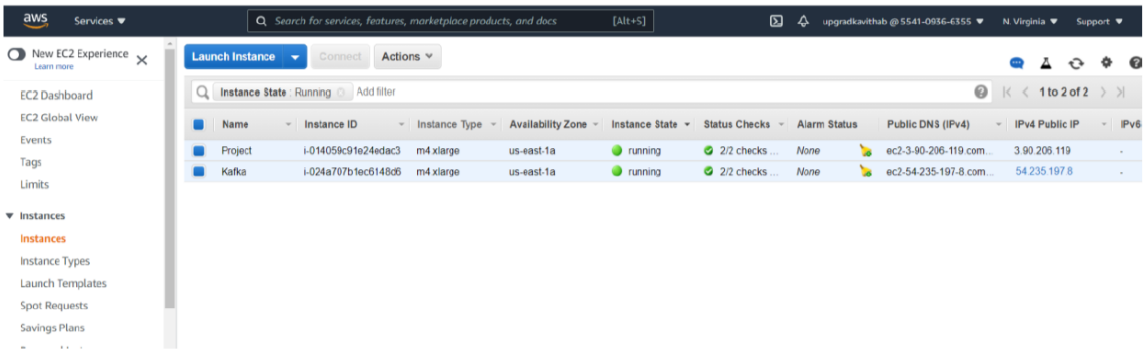
Creation and configuration of Kafka cluster

1.Create an kafka cluster with the help of ec2 instance.(Referred with PDF attached in the modules).Make the required configuration needed to run kafka instance.

2.Create a cdh instance using ec2 instance. .(Referred with PDF attached in the modules). Make the required configuration needed to run CDH instance



Steps to be done in before to create kafka Topics:

1.From ec2-user, navigate to [ec2-user@ip-172-31-27-4 kafka_2.12-2.3.0] by using **cd** command

2.Start the zookeeper by using the following command:

bin/zookeeper-server-start.sh config/zookeeper.properties

3.Start the kafka Server using the below command:

bin/kafka-server-start.sh config/server.properties

STATEMENT TO CREATE TOPICS

To create topic in kafka server ,the command used is

bin/kafka-topics.sh --create --bootstrap-server localhost:9092 --replication-factor 1 --partitions 1 --topic PatientInformation

STATEMENT TO LIST TOPICS

To list the topics in kafka server ,the command used is :

bin/kafka-topics.sh --list --bootstrap-server localhost:9092

STATEMENTS TO RUN PRODUCER AND CONSUMER PYTHON PROGRAMS

To run the Producer program ,use the below command ,

python prodcerapp.py

Note: For producer application,Kindly refer the kafka_Produce_patient_vitals.py

To run the consumer program ,use the below command ,

Patient_Consumer.py

export SPARK_KAFKA_VERSION=0.10

spark2-submit --jars spark-sql-kafka-0-10_2.11-2.3.0.jar Patient_Consumer.py

STATEMENTS TO CREATE EXTERNAL HIVE TABLE TO VIEW THRESHOLD DATA

1.Create a database using below command:

Create database if not exists Patient_db;

2.Use the database

Use patient_db;

3.Create a external hive table to create a threshold table.

Syntax :

create external table if not exists ThresholdValueTable(

Key int,

Attribute varchar(20),

Low_Age_Limit int,

High_Age_Limit int,

Low_range_value int,

High_range_value int,

Alert_Flag string,

Alert_Message string)

row format delimited fields terminated by "|"

lines terminated by "\n" stored as textfile;

4.Create a text file in hdfs in order to load the data,use the below command to create a text file in hdfs.

Vi threshold.txt

Enter the below data

1|heartbeat|0|40|0|69|1|Lowerheartratethannormal

2|heartbeat|0|40|70|78|0|Normal

3|heartbeat|0|40|79|9999|1|Higherheartratethannormal

4|bp|0|40|0|160|1|Lowerbpthannormal

5|bp|0|40|161|220|0|Normal

6|bp|0|40|221|9999|1|Higherbpthannormal

7|heartBeat|41|100|0|65|1|LowHeartRatethanNormal

8|heartBeat|41|100|66|73|0|Normal

9|heartBeat|41|100|74|9999|1|HigherHeartRatethanNormal

10|bp|41|100|0|150|1|LowBPthanNormal

11|bp|41|100|151|180|0|Normal

12|bp|41|100|181|9999|1|HigherBPthanNormal

5.Save the file by clicking **escape** key and save the file using **:wq!** 5.Load the data from text file into the created external table

```
hive> create external table if not exists ThresholdValueTable(
    > Key  int,
    > Attribute varchar(20),
    > Low_Age_Limit int,
    > High_Age_Limit int,
    > Low_range_value int,
    > High_range_value int,
    > Alert_Flag string,
    > Alert_Message string )
    > row format delimited fields terminated by  "|"
    > lines terminated by "\n" stored as textfile;
OK
Time taken: 0.088 seconds
hive> load data local inpath 'threshold.txt' into table ThresholdValueTable;
Loading data to table patient_db.thresholdvaluetable
Table patient_db.thresholdvaluetable stats: [numFiles=1, totalSize=486]
OK
Time taken: 1.153 seconds
hive> select* from thresholdvaluetable;
OK
1       heartbeat       0       40      0       69      1       Lowerheartratethannormal
2       heartbeat       0       40      70      78      0       Normal
3       heartbeat       0       40      79      9999    1       Higherheartratethannormal
4       bp      0       40      0       160     1       Lowerbpthannormal
5       bp      0       40      161     220     0       Normal
6       bp      0       40      221     9999    1       Higherbpthannormal
7       heartBeat       41      100     0       65      1       LowHeartRatethanNormal
8       heartBeat       41      100     66      73      0       Normal
9       heartBeat       41      100     74      9999    1       HigherHeartRatethanNormal
10      bp      41      100     0       150     1       LowBPthanNormal
11      bp      41      100     151     180     0       Normal
12      bp      41      100     181     9999    1       HigherBPthanNormal
```

STATEMENTS TO CREATE HBASE TABLE ON HIVE

1. From ec2-user, use the below command to go to hbase shell

   Hbase shell

1. 2. **Create hbase table using below command**:

create 'ThresholdTable', 'Attribute', 'Limit','Alert'

2.Describe the table column field using below command

describe 'ThresholdTable'

3.Insert the threshold data using put command for 12 records.

put 'ThresholdTable','row1','Limit:Low_age_limit','value=0'

put 'ThresholdTable','row1','Limit:High_age_limit','value=40'

put 'ThresholdTable','row1','Limit:Low_value','value=0'

put 'ThresholdTable','row1','Limit:High_value','value=69'

put 'ThresholdTable','row2','Limit:Low_age_limit','value=0'

put 'ThresholdTable','row2','Limit:High_age_limit','value=40'

put 'ThresholdTable','row2','Limit:Low_value','value=70'

put 'ThresholdTable','row2','Limit:High_value','value=78'

put 'ThresholdTable','row3','Limit:Low_age_limit','value=0'

put 'ThresholdTable','row3','Limit:High_age_limit','value=40'

put 'ThresholdTable','row3','Limit:Low_value','value=79'

put 'ThresholdTable','row3','Limit:High_value','value=9999'

put 'ThresholdTable','row4','Limit:Low_age_limit','value=0'

put 'ThresholdTable','row4','Limit:High_age_limit','value=40'

put 'ThresholdTable','row4','Limit:Low_value','value=0'

put 'ThresholdTable','row4','Limit:High_value','value=160'

```
put 'ThresholdTable','row5','Limit:Low_age_limit','value=0'
put 'ThresholdTable','row5','Limit:High_age_limit','value=40'
put 'ThresholdTable','row5','Limit:Low_value','value=161'
put 'ThresholdTable','row5','Limit:High_value','value=220'

put 'ThresholdTable','row6','Limit:Low_age_limit','value=0'
put 'ThresholdTable','row6','Limit:High_age_limit','value=40'
put 'ThresholdTable','row6','Limit:Low_value','value=221'
put 'ThresholdTable','row6','Limit:High_value','value=9999'

put 'ThresholdTable','row7','Limit:Low_age_limit','value=41'
put 'ThresholdTable','row7','Limit:High_age_limit','value=100'
put 'ThresholdTable','row7','Limit:Low_value','value=0'
put 'ThresholdTable','row7','Limit:High_value','value=65'

put 'ThresholdTable','row8','Limit:Low_age_limit','value=41'
put 'ThresholdTable','row8','Limit:High_age_limit','value=100'
put 'ThresholdTable','row8','Limit:Low_value','value=66'
put 'ThresholdTable','row8','Limit:High_value','value=73'

put 'ThresholdTable','row9','Limit:Low_age_limit','value=41'
put 'ThresholdTable','row9','Limit:High_age_limit','value=100'
put 'ThresholdTable','row9','Limit:Low_value','value=74'
put 'ThresholdTable','row9','Limit:High_value','value=9999'
```

```
put 'ThresholdTable','row10','Limit:Low_age_limit','value=41'
put 'ThresholdTable','row10','Limit:High_age_limit','value=100'
put 'ThresholdTable','row10','Limit:Low_value','value=0'
put 'ThresholdTable','row10','Limit:High_value','value=150'

put 'ThresholdTable','row11','Limit:Low_age_limit','value=41'
put 'ThresholdTable','row11','Limit:High_age_limit','value=100'
put 'ThresholdTable','row11','Limit:Low_value','value=151'
put 'ThresholdTable','row11','Limit:High_value','value=180'

put 'ThresholdTable','row12','Limit:Low_age_limit','value=41'
put 'ThresholdTable','row12','Limit:High_age_limit','value=100'
put 'ThresholdTable','row12','Limit:Low_value','value=181'
put 'ThresholdTable','row12','Limit:High_value','value=9999'

put 'ThresholdTable','row1','Alert:Alert_Message','value=Low Heart Rate than Normal'
put 'ThresholdTable','row2','Alert:Alert_Message','value=Normal'
put 'ThresholdTable','row3','Alert:Alert_Message','value=Higher Heart Rate than Normal'
put 'ThresholdTable','row4','Alert:Alert_Message','value=Low BP than Normal'
put 'ThresholdTable','row5','Alert:Alert_Message','value=Normal'
put 'ThresholdTable','row6','Alert:Alert_Message','value=Higher BP than Normal'
put 'ThresholdTable','row7','Alert:Alert_Message','value=Low Heart Rate than Normal'
put 'ThresholdTable','row8','Alert:Alert_Message','value=Normal'
put 'ThresholdTable','row9','Alert:Alert_Message','value=Higher Heart Rate than Normal'
put 'ThresholdTable','row10','Alert:Alert_Message','value=Low BP than Normal'
```

put 'ThresholdTable','row11','Alert:Alert_Message','value=Normal'

put 'ThresholdTable','row12','Alert:Alert_Message','value=Higher BP than Normal'

put 'ThresholdTable','row1','Attribute:Attribute','value=Heartbeat'

put 'ThresholdTable','row2','Attribute:Attribute','value=Heartbeat'

put 'ThresholdTable','row3','Attribute:Attribute','value=Heartbeat'

put 'ThresholdTable','row4','Attribute:Attribute','value=BP'

put 'ThresholdTable','row5','Attribute:Attribute','value=BP'

put 'ThresholdTable','row6','Attribute:Attribute','value=BP'

put 'ThresholdTable','row7','Attribute:Attribute','value=Heartbeat'

put 'ThresholdTable','row8','Attribute:Attribute','value=Heartbeat'

put 'ThresholdTable','row9','Attribute:Attribute','value=Heartbeat'

put 'ThresholdTable','row10','Attribute:Attribute','value=BP'

put 'ThresholdTable','row11','Attribute:Attribute','value=BP'

put 'ThresholdTable','row12','Attribute:Attribute','value=BP'

put 'ThresholdTable','row1','Alert:Alert_Flag','value=1'

put 'ThresholdTable','row2','Alert:Alert_Flag','value=0'

put 'ThresholdTable','row3','Alert:Alert_Flag','value=1'

put 'ThresholdTable','row4','Alert:Alert_Flag','value=1'

put 'ThresholdTable','row5','Alert:Alert_Flag','value=0'

put 'ThresholdTable','row6','Alert:Alert_Flag','value=1'

put 'ThresholdTable','row7','Alert:Alert_Flag','value=1'

put 'ThresholdTable','row8','Alert:Alert_Flag','value=0'

put 'ThresholdTable','row9','Alert:Alert_Flag','value=1'

put 'ThresholdTable','row10','Alert:Alert_Flag','value=1'

put 'ThresholdTable','row11','Alert:Alert_Flag','value=0'

put 'ThresholdTable','row12','Alert:Alert_Flag','value=1'

4.To view the table,use the below command

Scan ‘ThresholdTable’

Screenshots:

```
hbase(main):107:0> scan 'ThresholdTable'
ROW                                     COLUMN+CELL
 row1                                   column=Alert:Alert_Flag, timestamp=1636803743544, value=value=1
 row1                                   column=Alert:Alert_Message, timestamp=1636803963284, value=value=Low Heart Rate than Normal
 row1                                   column=Attribute:Attribute, timestamp=1636803743376, value=value=Heartbeat
 row1                                   column=Limit:High_age_limit, timestamp=1636803742467, value=value=40
 row1                                   column=Limit:High_value, timestamp=1636803742520, value=value=69
 row1                                   column=Limit:Low_age_limit, timestamp=1636803742423, value=value=0
 row1                                   column=Limit:Low_value, timestamp=1636803742500, value=value=0
 row10                                  column=Alert:Alert_Flag, timestamp=1636803743646, value=value=1
 row10                                  column=Alert:Alert_Message, timestamp=1636803963422, value=value=Low BP than Normal
 row10                                  column=Attribute:Attribute, timestamp=1636803743495, value=value=BP
 row10                                  column=Limit:High_age_limit, timestamp=1636803743127, value=value=100
 row10                                  column=Limit:High_value, timestamp=1636803743152, value=value=150
 row10                                  column=Limit:Low_age_limit, timestamp=1636803743116, value=value=41
 row10                                  column=Limit:Low_value, timestamp=1636803743140, value=value=0
 row11                                  column=Alert:Alert_Flag, timestamp=1636803743657, value=value=0
 row11                                  column=Alert:Alert_Message, timestamp=1636803963436, value=value=Normal
 row11                                  column=Attribute:Attribute, timestamp=1636803743507, value=value=BP
 row11                                  column=Limit:High_age_limit, timestamp=1636803743179, value=value=100
 row11                                  column=Limit:High_value, timestamp=1636803743203, value=value=180
 row11                                  column=Limit:Low_age_limit, timestamp=1636803743169, value=value=41
 row11                                  column=Limit:Low_value, timestamp=1636803743190, value=value=151
 row12                                  column=Alert:Alert_Flag, timestamp=1636803743667, value=value=1
 row12                                  column=Alert:Alert_Message, timestamp=1636803966435, value=value=Higher BP than Normal
 row12                                  column=Attribute:Attribute, timestamp=1636803743519, value=value=BP
 row12                                  column=Limit:High_age_limit, timestamp=1636803743293, value=value=100
 row12                                  column=Limit:High_value, timestamp=1636803743340, value=value=9999
 row12                                  column=Limit:Low_age_limit, timestamp=1636803743251, value=value=41
 row12                                  column=Limit:Low_value, timestamp=1636803743324, value=value=181
 row2                                   column=Alert:Alert_Flag, timestamp=1636803743554, value=value=0
 row2                                   column=Alert:Alert_Message, timestamp=1636803963306, value=value=Normal
 row2                                   column=Attribute:Attribute, timestamp=1636803743403, value=value=Heartbeat
 row2                                   column=Limit:High_age_limit, timestamp=1636803742561, value=value=40
 row2                                   column=Limit:High_value, timestamp=1636803742594, value=value=78
 row2                                   column=Limit:Low_age_limit, timestamp=1636803742547, value=value=0
 row2                                   column=Limit:Low_value, timestamp=1636803742575, value=value=70
 row3                                   column=Alert:Alert_Flag, timestamp=1636803743563, value=value=1
 row3                                   column=Alert:Alert_Message, timestamp=1636803963331, value=value=Higher Heart Rate than Normal
 row3                                   column=Attribute:Attribute, timestamp=1636803743416, value=value=Heartbeat
 row3                                   column=Limit:High_age_limit, timestamp=1636803742634, value=value=40
 row3                                   column=Limit:High_value, timestamp=1636803742685, value=value=9999
 row3                                   column=Limit:Low_age_limit, timestamp=1636803742619, value=value=0
 row3                                   column=Limit:Low_value, timestamp=1636803742657, value=value=79
 row4                                   column=Alert:Alert_Flag, timestamp=1636803743574, value=value=1
 row4                                   column=Alert:Alert_Message, timestamp=1636803963349, value=value=Low BP than Normal
```

```
 row3                                   column=Attribute:Attribute, timestamp=1636803743416, value=value=Heartbeat
 row3                                   column=Limit:High_age_limit, timestamp=1636803742634, value=value=40
 row3                                   column=Limit:High_value, timestamp=1636803742685, value=value=9999
 row3                                   column=Limit:Low_age_limit, timestamp=1636803742619, value=value=0
 row3                                   column=Limit:Low_value, timestamp=1636803742657, value=value=79
 row4                                   column=Alert:Alert_Flag, timestamp=1636803743574, value=value=1
 row4                                   column=Alert:Alert_Message, timestamp=1636803963349, value=value=Low BP than Normal
 row4                                   column=Attribute:Attribute, timestamp=1636803743428, value=value=BP
 row4                                   column=Limit:High_age_limit, timestamp=1636803742734, value=value=40
 row4                                   column=Limit:High_value, timestamp=1636803742770, value=value=160
 row4                                   column=Limit:Low_age_limit, timestamp=1636803742708, value=value=0
 row4                                   column=Limit:Low_value, timestamp=1636803742748, value=value=0
 row5                                   column=Alert:Alert_Flag, timestamp=1636803743584, value=value=0
 row5                                   column=Alert:Alert_Message, timestamp=1636803963363, value=value=Normal
 row5                                   column=Attribute:Attribute, timestamp=1636803743440, value=value=BP
 row5                                   column=Limit:High_age_limit, timestamp=1636803742801, value=value=40
 row5                                   column=Limit:High_value, timestamp=1636803742828, value=value=220
 row5                                   column=Limit:Low_age_limit, timestamp=1636803742789, value=value=0
 row5                                   column=Limit:Low_value, timestamp=1636803742815, value=value=161
 row6                                   column=Alert:Alert_Flag, timestamp=1636803743594, value=value=1
 row6                                   column=Alert:Alert_Message, timestamp=1636803963377, value=value=Higher BP than Normal
 row6                                   column=Attribute:Attribute, timestamp=1636803743452, value=value=BP
 row6                                   column=Limit:High_age_limit, timestamp=1636803742861, value=value=40
 row6                                   column=Limit:High_value, timestamp=1636803742903, value=value=9999
 row6                                   column=Limit:Low_age_limit, timestamp=1636803742845, value=value=0
 row6                                   column=Limit:Low_value, timestamp=1636803742880, value=value=221
 row7                                   column=Alert:Alert_Flag, timestamp=1636803743604, value=value=1
 row7                                   column=Alert:Alert_Message, timestamp=1636803963389, value=value=Low Heart Rate than Normal
 row7                                   column=Attribute:Attribute, timestamp=1636803743464, value=value=Heartbeat
 row7                                   column=Limit:High_age_limit, timestamp=1636803742938, value=value=100
 row7                                   column=Limit:High_value, timestamp=1636803742962, value=value=65
 row7                                   column=Limit:Low_age_limit, timestamp=1636803742925, value=value=41
 row7                                   column=Limit:Low_value, timestamp=1636803742951, value=value=0
 row8                                   column=Alert:Alert_Flag, timestamp=1636803743616, value=value=0
 row8                                   column=Alert:Alert_Message, timestamp=1636803963398, value=value=Normal
 row8                                   column=Attribute:Attribute, timestamp=1636803743475, value=value=Heartbeat
 row8                                   column=Limit:High_age_limit, timestamp=1636803743003, value=value=100
 row8                                   column=Limit:High_value, timestamp=1636803743036, value=value=73
 row8                                   column=Limit:Low_age_limit, timestamp=1636803742984, value=value=41
 row8                                   column=Limit:Low_value, timestamp=1636803743022, value=value=66
 row9                                   column=Alert:Alert_Flag, timestamp=1636803743635, value=value=1
 row9                                   column=Alert:Alert_Message, timestamp=1636803963412, value=value=Higher Heart Rate than Normal
 row9                                   column=Attribute:Attribute, timestamp=1636803743485, value=value=Heartbeat
 row9                                   column=Limit:High_age_limit, timestamp=1636803743063, value=value=100
 row9                                   column=Limit:High_value, timestamp=1636803743093, value=value=9999
 row9                                   column=Limit:Low_age_limit, timestamp=1636803743053, value=value=41
 row9                                   column=Limit:Low_value, timestamp=1636803743076, value=value=74
12 row(s) in 0.0320 seconds
```

STATEMENTS TO CREATE EXTERNAL HIVE TABLE FOR PATIENTS VITAL INFORMATION

1. **Creating externaltal table for patient vital information**

```
hive> CREATE EXTERNAL TABLE PatientInfoTable (
    >  BP  String,
    >  CustomerId String,
    >  Heartbeat String,
    >  timestamp String)
    >  STORED AS PARQUET;
```

2.Loading Data from parquet files into hive table.

```
load data inpath 'PatientInformation' into table PatientInfoTable;
```

```
Logging initialized using configuration in jar:file:/opt/cloudera/parcels/CDH-5.15.1-1.cdh5.15.1.p0.4/jars/hive-common-1.1.0-cdh5.15.1.jar!/hive-log4j.properties
WARNING: Hive CLI is deprecated and migration to Beeline is recommended.
hive> CREATE EXTERNAL TABLE PatientInfoTable (
    >     BP  String,
    >     CustomerId String,
    >     Heartbeat String,
    >     timestamp String)
    >     STORED AS PARQUET;
OK
Time taken: 1.898 seconds
hive> load data local  inpath 'PatientInformation' into table PatientInfoTable;
FAILED: SemanticException Line 1:24 Invalid path ''PatientInformation'': No files matching path file:/var/lib/hadoop-hdfs/PatientInformation
hive> load data inpath 'PatientInformation' into table PatientInfoTable;
Loading data to table default.patientinfocable
Table default.patientinfocable stats: [numFiles=3, numRows=0, totalSize=4061, rawDataSize=0]
OK
Time taken: 0.355 seconds
hive> select * from  PatientInfoTable;
OK
```

3.To view the data in PatientInfoTable. Execute the below command

hive> select * from PatientInfoTable;

Data Retrieved Screenshot:

```
158     3       66      2021-11-13 13:46:00.318
176     4       71      2021-11-13 13:46:01.32
168     5       67      2021-11-13 13:46:02.322
184     1       73      2021-11-13 13:46:03.324
160     2       73      2021-11-13 13:46:04.327
156     3       67      2021-11-13 13:46:05.328
154     4       66      2021-11-13 13:46:06.33
172     5       66      2021-11-13 13:46:07.331
212     1       75      2021-11-13 13:46:08.333
165     2       67      2021-11-13 13:46:09.335
154     3       68      2021-11-13 13:46:10.337
178     4       71      2021-11-13 13:46:11.339
177     5       69      2021-11-13 13:46:12.342
174     1       73      2021-11-13 13:46:13.344
178     2       69      2021-11-13 13:46:14.347
176     3       70      2021-11-13 13:46:15.348
161     4       73      2021-11-13 13:46:16.35
167     5       70      2021-11-13 13:46:17.352
191     1       71      2021-11-13 13:46:18.354
180     2       67      2021-11-13 13:46:19.356
159     3       69      2021-11-13 13:46:20.358
175     4       68      2021-11-13 13:46:21.36
162     5       66      2021-11-13 13:46:22.362
172     1       70      2021-11-13 13:46:23.363
179     2       69      2021-11-13 13:46:24.365
152     3       68      2021-11-13 13:46:25.367
163     4       72      2021-11-13 13:46:26.369
173     5       70      2021-11-13 13:46:27.371
168     1       78      2021-11-13 13:46:28.372
162     2       67      2021-11-13 13:46:29.374
177     3       66      2021-11-13 13:46:30.376
168     4       73      2021-11-13 13:46:31.378
155     5       71      2021-11-13 13:46:32.38
167     1       76      2021-11-13 13:46:33.382
178     2       73      2021-11-13 13:46:34.384
```

STATEMENTS TO RETRIEVE THE PATIENT'S CONTACT DETAILS USING SQOOP

1.Extract patient contact details by executing the below sqoop command

```
sqoop import \

--connect jdbc:mysql://upgraddetest.cyaielc9bmnf.us-east-1.rds.amazonaws.com/testdatabase \

--username student \

--password STUDENT123 \

--table patients_information \

--direct \

-m 1 \

--hive-import \

--create-hive-table \

--hive-table PatientContactInfo \

--fields-terminated-by ';'
```

SNAPSHOT OF DATA COLLECTED

```
[hdfs@ip-10-0-0-227 ~]$ sqoop import \
> --connect jdbc:mysql://upgraddetest.cyaielc9bmnf.us-east-1.rds.amazonaws.com/testdatabase \
> --username student \
> --password STUDENT123 \
> --table patients_information \
> --direct \
> -m 1 \
> --hive-import \
> --create-hive-table \
> --hive-table PatientContactInfo \
> --fields-terminated-by ';'
Warning: /opt/cloudera/parcels/CDH-5.15.1-1.cdh5.15.1.p0.4/bin/../lib/sqoop/../accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
21/11/13 09:50:25 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6-cdh5.15.1
21/11/13 09:50:25 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
21/11/13 09:50:25 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
21/11/13 09:50:25 INFO tool.CodeGenTool: Beginning code generation
21/11/13 09:50:25 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `patients_information` AS t LIMIT 1
21/11/13 09:50:25 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `patients_information` AS t LIMIT 1
21/11/13 09:50:25 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /opt/cloudera/parcels/CDH/lib/hadoop-mapreduce
Note: /tmp/sqoop-hdfs/compile/968194ca6a5a4cb6300973df91d349cd/patients_information.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
21/11/13 09:50:28 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-hdfs/compile/968194ca6a5a4cb6300973df91d349cd/patients_information.jar
21/11/13 09:50:28 INFO manager.DirectMySQLManager: Beginning mysqldump fast path import
21/11/13 09:50:28 INFO mapreduce.ImportJobBase: Beginning import of patients_information
21/11/13 09:50:28 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar
21/11/13 09:50:29 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps
21/11/13 09:50:30 INFO client.RMProxy: Connecting to ResourceManager at ip-10-0-0-227.ec2.internal/10.0.0.227:8032
21/11/13 09:50:45 INFO db.DBInputFormat: Using read commited transaction isolation
21/11/13 09:50:45 INFO mapreduce.JobSubmitter: number of splits:1
21/11/13 09:50:46 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1636795331014_0004
21/11/13 09:50:46 INFO impl.YarnClientImpl: Submitted application application_1636795331014_0004
21/11/13 09:50:46 INFO mapreduce.Job: The url to track the job: http://ip-10-0-0-227.ec2.internal:8088/proxy/application_1636795331014_0004/
21/11/13 09:50:46 INFO mapreduce.Job: Running job: job_1636795331014_0004
21/11/13 09:50:53 INFO mapreduce.Job: Job job_1636795331014_0004 running in uber mode : false
21/11/13 09:50:53 INFO mapreduce.Job:  map 0% reduce 0%
21/11/13 09:51:00 INFO mapreduce.Job:  map 100% reduce 0%
```

2.Extarcted data in hive table After executing sqoop command

```
hive> select * from patientcontactinfo
    > ;
OK
1       Alex S  XDC test Address        8982739282      1       23      NULL
2       Sammy A New Building Address    2382739282      2       45      NULL
3       Karan C Aws Address     8923739282      3       56      NULL
4       Dara M  India Address   2182739282      4       67      NULL
5       Pam     ABC test Address        4982739282      5       72      NULL
Time taken: 0.096 seconds, Fetched: 5 row(s)
hive>
```

STATEMENTS TO READ DATA FROM HDFS AND COMPARE WITH Hbase TABLE

Run the spark streaming application named as kafka_spark_generate_alerts.py,use the below command to execute the program

Python kafka_spark_generate_alerts.py

STATEMENTS TO EXECUTE THE CONSUMER APPLICATION FOR SENDING ALERTS

1.From ec2-user, navigate to [ec2-user@ip-172-31-27-4 kafka_2.12-2.3.0] by using **cd** command

2.Start the zookeeper by using the following command:

bin/zookeeper-server-start.sh config/zookeeper.properties

3.Start the kafka Server using the below command:

bin/kafka-server-start.sh config/server.properties

STATEMENT TO CREATE TOPICS

To create topic in kafka server ,the command used is

bin/kafka-topics.sh --create --bootstrap-server localhost:9092 --replication-factor 1 --partitions 1 --topic PatientHealthNotification

To list the topics in kafka server ,the command used is :

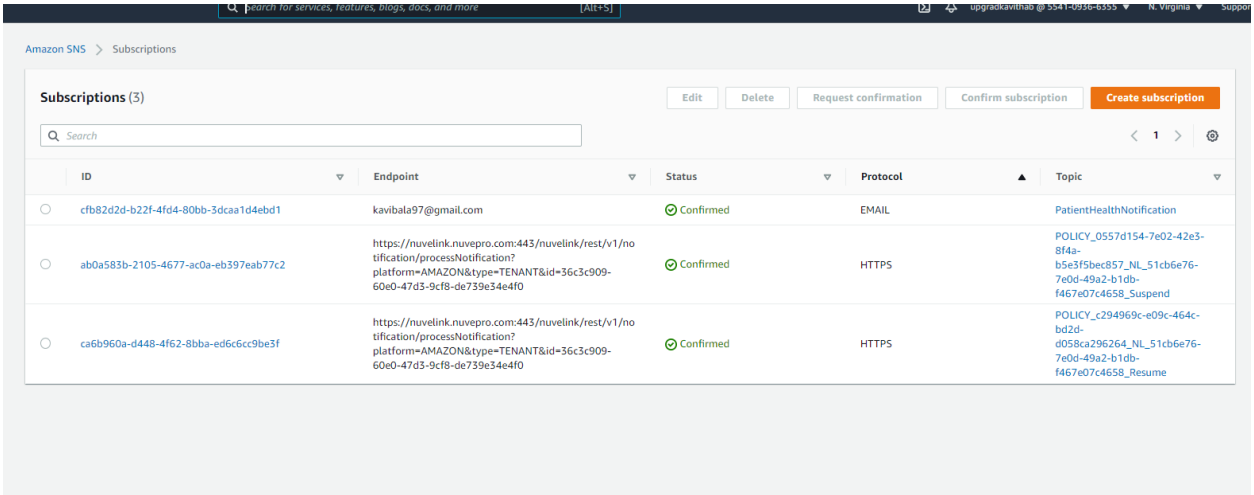
bin/kafka-topics.sh --list --bootstrap-server localhost:9092

Run the below command to start the producer .

bin/kafka-console-producer.sh --broker-list localhost:9092 –topic PatientHealthNotification

Open the another terminal in putty , run the consumer program named kafka_consume_alerts.py using below command

Python kafka_consume_alerts.py



3.Confirm the subscription in the email sent to the mentioned email id.