1. Create a schema based on the given dataset

```
hive> CREATE TABLE IF NOT EXISTS AgentPerformance
    > (
    > `SL No` INT,
    > `Date` STRING,
    > `Agent Name` STRING,
    > `Total Chats` INT,
    > `Average Response Time` STRING,
    > `Average Resolution Time` STRING,
    > `Average Rating` FLOAT,
    > `Total Feedback` INT
    > )
    > row format delimited
    > fields terminated by ','
    > tblproperties("skip.header.line.count" = "1");
OK
Time taken: 0.837 seconds
hive> CREATE TABLE IF NOT EXISTS AgentLogingReport
    > (
    > `SL No` INT,
    > Agent STRING,
    > `Date` STRING,
    > `Login Time` STRING,
    > `Logout Time` STRING,
    > Duration STRING
    > )
    > row format delimited
    > fields terminated by ','
    > tblproperties("skip.header.line.count" = "1");
OK
Time taken: 0.239 seconds
hive>
```

2. Dump the data inside the hdfs in the given schema location.

load data local inpath 'file:///home/cloudera/Hive_class/AgentPerformance.csv' into table AgentPerformance;

load data local inpath 'file:///home/cloudera/Hive_class/AgentLogingReport.csv' into table AgentLogingReport;

```
hive> load data local inpath 'file:///home/cloudera/Hive_class/AgentPerformance.csv' into table AgentPerformance;
Loading data to table hive_class_b1.agentperformance
Table hive_class_b1.agentperformance stats: [numFiles=1, totalSize=109853]
OK
Time taken: 1.042 seconds
hive> load data local inpath 'file:///home/cloudera/Hive_class/AgentLogingReport.csv' into table AgentLogingReport;
Loading data to table hive_class_b1.agentlogingreport
Table hive_class_b1.agentlogingreport stats: [numFiles=1, totalSize=55351]
OK
Time taken: 0.921 seconds
hive>
```

3. List of all agents' names.

SELECT DISTINCT `Agent Name` FROM AgentPerformance;

```
hive> SELECT DISTINCT `Agent Name` FROM AgentPerformance;
Query ID = cloudera_20220930015252_0afab604-30c8-4298-a972-c988b5b64f8f
Total jobs = 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_1664521724151_0003, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1664521724151_0003/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1664521724151_0003
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-09-30 01:52:50,397 Stage-1 map = 0%,  reduce = 0%
2022-09-30 01:53:00,743 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 2.38 sec
2022-09-30 01:53:10,775 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 5.14 sec
MapReduce Total cumulative CPU time: 5 seconds 140 msec
Ended Job = job_1664521724151_0003
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 5.14 sec   HDFS Read: 118523 HDFS Write: 867 SUCCESS
Total MapReduce CPU Time Spent: 5 seconds 140 msec
OK
agent name
Abhishek
Aditya
Aditya Shinde
Aditya_iot
Amersh
Ameya Jain
Anirudh
Ankit Sharma
Ankitjha
Anurag Tiwari
Aravind
Ashad Nasim
Ashish
Ayushi Mishra
Bharath
Boktiar Ahmed Bappy
Chaitra K Hiremath
Deepranjan Gupta
Dibyanshu
Harikrishnan Shaji
Hitesh Choudhary
Hrisikesh Neogi
Hyder Abbas
Ineuron Intelligence
Ishawant Kumar
Jawala Prakash
```

```
Hrisikesh Neogi
Hyder Abbas
Ineuron Intelligence
Ishawant Kumar
Jawala Prakash
Jayant Kumar
Jaydeep Dixit
Khushboo Priya
Madhulika G
Mahak
Mahesh Sarade
Maitry
Maneesh
Manjunatha A
Mithun S
Mukesh
Mukesh Rao
Muskan Garg
Nandani Gupta
Nishtha Jain
Nitin M
Prabir Kumar Satapathy
Prateek _iot
Prerna Singh
Rishav Dash
Rohan
Saif Khan
Saikumarreddy N
Samprit
Sandipan Saha
Sanjeev Kumar
Sanjeevan
Saurabh Shukla
Shiva Srivastava
Shivan K
Shivan_S
Shivananda Sonwane
Shubham Sharma
Sowmiya Sivakumar
Spuri
Sudhanshu Kumar
Suraj S Bilgi
Swati
Tarun
Uday Mishra
Vasanth P
Vivek
Wasim
Zeeshan
Time taken: 35.254 seconds, Fetched: 70 row(s)
```

4. Find out agent average rating.

SELECT `Agent Name`, AVG(`Average Rating`) AS `Agent Average Rating` FROM AgentPerformance GROUP BY `Agent Name`;

```
hive> SELECT `Agent Name`, AVG(`Average Rating`) AS `Agent Average Rating` FROM AgentPerformance GROUP BY `Agent Name`;
Query ID = cloudera_20220930015656_e24cc5ae-9307-48ac-b3e5-bdc526ac3956
Total jobs = 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_1664521724151_0004, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1664521724151_0004/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1664521724151_0004
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-09-30 01:56:21,851 Stage-1 map = 0%,  reduce = 0%
2022-09-30 01:56:33,184 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 2.75 sec
2022-09-30 01:56:43,002 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 6.1 sec
MapReduce Total cumulative CPU time: 6 seconds 100 msec
Ended Job = job_1664521724151_0004
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 6.1 sec   HDFS Read: 119769 HDFS Write: 1904 SUCCESS
Total MapReduce CPU Time Spent: 6 seconds 100 msec
OK
agent name      agent average rating
Abhishek        0.0
Aditya  0.0
Aditya Shinde   1.8003333409627278
Aditya_iot      2.3453333377838135
Amersh  0.0
Ameya Jain      2.21966667175293
Anirudh         0.6449999968210857
Ankit Sharma    0.0
Ankitjha        0.26666666666666666
Anurag Tiwari   0.18333333333333332
Aravind         2.1813333511352537
Ashad Nasim     0.16666666666666666
Ashish  0.0
Ayushi Mishra   3.481999969482422
Bharath         2.9836666584014893
Boktiar Ahmed Bappy     3.567999982833862
Chaitra K Hiremath      0.8646666600267294
Deepranjan Gupta        2.886666695276896
Dibyanshu       0.0
Harikrishnan Shaji      2.6396666526794434
Hitesh Choudhary        0.0
Hrisikesh Neogi 3.1363333304723104
Hyder Abbas     0.0
Ineuron Intelligence    0.0
Ishawant Kumar  3.543333347638448
Jawala Prakash  3.472000018755595
```

```
hive> SELECT `Agent Name`, AVG(`Average Rating`) AS `Agent Average Rating` FROM AgentPerformance GROUP BY `Agent Name`;
Query ID = cloudera_20220930015656_e24cc5ae-9307-48ac-b3e5-bdc526ac3956
Total jobs = 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_1664521724151_0004, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1664521724151_0004/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1664521724151_0004
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-09-30 01:56:21,851 Stage-1 map = 0%,  reduce = 0%
2022-09-30 01:56:33,184 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 2.75 sec
2022-09-30 01:56:43,002 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 6.1 sec
MapReduce Total cumulative CPU time: 6 seconds 100 msec
Ended Job = job_1664521724151_0004
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 6.1 sec   HDFS Read: 119769 HDFS Write: 1904 SUCCESS
Total MapReduce CPU Time Spent: 6 seconds 100 msec
OK
agent name      agent average rating
Abhishek        0.0
Aditya  0.0
Aditya Shinde   1.8003333409627278
Aditya_iot      2.3453333377838135
Amersh  0.0
Ameya Jain      2.21966667175293
Anirudh         0.6449999968210857
Ankit Sharma    0.0
Ankitjha        0.26666666666666666
Anurag Tiwari   0.18333333333333332
Aravind         2.1813333511352537
Ashad Nasim     0.16666666666666666
Ashish  0.0
Ayushi Mishra   3.481999969482422
Bharath         2.9836666584014893
Boktiar Ahmed Bappy     3.567999982833862
Chaitra K Hiremath      0.8646666606267294
Deepranjan Gupta        2.886666695276896
Dibyanshu       0.0
Harikrishnan Shaji      2.6396666526794434
Hitesh Choudhary        0.0
Hrisikesh Neogi 3.1363333304723104
Hyder Abbas     0.0
Ineuron Intelligence    0.0
Ishawant Kumar  3.543333347638448
Jawala Prakash  3.472000018755595
```

5. Total working days for each agents

select agent, count(distinct(date)) as Logged from agentlogingreport group by agent;

```
hive> select agent, count(distinct(date)) as Logged from agentlogingreport group by agent;
Query ID = cloudera_20221008092929_a36e8d6e-c5a1-4279-bdd2-70ae835a0292
Total jobs = 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_1665242270977_0003, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1665242270977_0003/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1665242270977_0003
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-08 09:30:02,145 Stage-1 map = 0%,  reduce = 0%
2022-10-08 09:30:13,132 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.98 sec
2022-10-08 09:30:23,387 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 4.93 sec
MapReduce Total cumulative CPU time: 4 seconds 930 msec
Ended Job = job_1665242270977_0003
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 4.93 sec   HDFS Read: 63940 HDFS Write: 742 SUCCESS
Total MapReduce CPU Time Spent: 4 seconds 930 msec
OK
agent   logged
Aditya Shinde   1
Aditya_iot      8
Amersh  2
Ameya Jain      7
Ankitjha        2
Anurag Tiwari   10
Aravind 7
Ayushi Mishra   9
Bharath 8
Boktiar Ahmed Bappy     9
Chaitra K Hiremath      7
Deepranjan Gupta        10
Dibyanshu       9
Harikrishnan Shaji      9
Hrisikesh Neogi 9
Hyder Abbas     2
Ineuron Intelligence    1
Ishawant Kumar  11
Jawala Prakash  9
Jaydeep Dixit   7
Khushboo Priya  8
Madhulika G     8
Mahesh Sarade   8
Maitry  5
Manjunatha A    7
Mithun S        8
```

6. Total query that each agent have taken.

SELECT `Agent Name`, SUM(`Total chats`) AS `Total query` FROM AgentPerformance GROUP BY `Agent Name`;

```
hive> SELECT `Agent Name`, SUM(`Total chats`) AS `Total query` FROM AgentPerformance GROUP BY `Agent Name`;
Query ID = cloudera_20220930032222_d1a4026c-c351-4b3f-a94c-4797b8fb2325
Total jobs = 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_1664521724151_0011, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1664521724151_0011/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1664521724151_0011
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-09-30 03:22:47,253 Stage-1 map = 0%,  reduce = 0%
2022-09-30 03:22:52,668 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.52 sec
2022-09-30 03:23:00,667 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 3.39 sec
MapReduce Total cumulative CPU time: 3 seconds 390 msec
Ended Job = job_1664521724151_0011
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 3.39 sec   HDFS Read: 118878 HDFS Write: 1088 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 390 msec
OK
agent name      total query
Abhishek        0
Aditya  0
Aditya Shinde   277
Aditya_iot      231
Amersh  0
Ameya Jain      322
Anirudh         81
Ankit Sharma    0
Ankitjha        5
Anurag Tiwari   4
Aravind         366
Ashad Nasim     18
Ashish  0
Ayushi Mishra   514
Bharath         369
Boktiar Ahmed Bappy     452
Chaitra K Hiremath      64
Deepranjan Gupta        493
Dibyanshu       1
Harikrishnan Shaji      381
Hitesh Choudhary        1
Hrisikesh Neogi 578
Hyder Abbas     0
Ineuron Intelligence    0
Ishawant Kumar  338
Jawala Prakash  439
```

```
Hyder Abbas     0
Ineuron Intelligence    0
Ishawant Kumar  338
Jawala Prakash  439
Jayant Kumar    127
Jaydeep Dixit   512
Khushboo Priya  446
Madhulika G     469
Mahak   7
Mahesh Sarade   364
Maitry  542
Maneesh         4
Manjunatha A    413
Mithun S        503
Mukesh  19
Mukesh Rao      5
Muskan Garg     56
Nandani Gupta   560
Nishtha Jain    373
Nitin M 0
Prabir Kumar Satapathy  299
Prateek _iot    190
Prerna Singh    401
Rishav Dash     409
Rohan   0
Saif Khan       0
Saikumarreddy N 364
Samprit         1
Sandipan Saha   30
Sanjeev Kumar   507
Sanjeevan       0
Saurabh Shukla  16
Shiva Srivastava        53
Shivan K        357
Shivan_S        7
Shivananda Sonwane      441
Shubham Sharma  510
Sowmiya Sivakumar       206
Spuri   0
Sudhanshu Kumar 2
Suraj S Bilgi   28
Swati   524
Tarun   22
Uday Mishra     0
Vasanth P       0
Vivek   44
Wasim   433
Zeeshan         542
Time taken: 23.199 seconds, Fetched: 70 row(s)
hive>
```

7. Total Feedback that each agent have received.

SELECT `Agent Name`, SUM(`Total Feedback`) AS `Total Feedback` FROM AgentPerformance GROUP BY `Agent Name`;

```
hive> SELECT `Agent Name`, SUM(`Total Feedback`) AS `Total Feedback` FROM AgentPerformance GROUP BY `Agent Name`;
Query ID = cloudera_20220930032525_d8a67a2f-9956-4070-8d73-c5f9a3a7c14c
Total jobs = 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_1664521724151_0012, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1664521724151_0012/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1664521724151_0012
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-09-30 03:26:06,586 Stage-1 map = 0%,  reduce = 0%
2022-09-30 03:26:12,894 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.26 sec
2022-09-30 03:26:19,330 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 3.07 sec
MapReduce Total cumulative CPU time: 3 seconds 70 msec
Ended Job = job_1664521724151_0012
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 3.07 sec   HDFS Read: 118884 HDFS Write: 1084 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 70 msec
OK
agent name      total feedback
Abhishek        0
Aditya  0
Aditya Shinde   153
Aditya_iot      131
Amersh  0
Ameya Jain      228
Anirudh         39
Ankit Sharma    0
Ankitjha        3
Anurag Tiwari   3
Aravind         233
Ashad Nasim     9
Ashish  0
Ayushi Mishra   329
Bharath         247
Boktiar Ahmed Bappy     311
Chaitra K Hiremath      37
Deepranjan Gupta        312
Dibyanshu       0
Harikrishnan Shaji      231
Hitesh Choudhary        0
Hrisikesh Neogi 367
Hyder Abbas     0
Ineuron Intelligence    0
Ishawant Kumar  202
Jawala Prakash  250
```

```
Hyder Abbas     0
Ineuron Intelligence    0
Ishawant Kumar  202
Jawala Prakash  250
Jayant Kumar    70
Jaydeep Dixit   305
Khushboo Priya  289
Madhulika G     281
Mahak   5
Mahesh Sarade   216
Maitry  347
Maneesh         3
Manjunatha A    254
Mithun S        364
Mukesh  17
Mukesh Rao      5
Muskan Garg     37
Nandani Gupta   308
Nishtha Jain    257
Nitin M 0
Prabir Kumar Satapathy  222
Prateek _iot    107
Prerna Singh    235
Rishav Dash     264
Rohan   0
Saif Khan       0
Saikumarreddy N 290
Samprit         0
Sandipan Saha   18
Sanjeev Kumar   311
Sanjeevan       0
Saurabh Shukla  8
Shiva Srivastava        46
Shivan K        243
Shivan_S        4
Shivananda Sonwane      263
Shubham Sharma  300
Sowmiya Sivakumar       141
Spuri   0
Sudhanshu Kumar 2
Suraj S Bilgi   15
Swati   302
Tarun   6
Uday Mishra     0
Vasanth P       0
Vivek   20
Wasim   284
Zeeshan         335
Time taken: 23.27 seconds, Fetched: 70 row(s)
hive>
```

8. Agent name who have average rating between 3.5 to 4.

SELECT DISTINCT `Agent Name` FROM AgentPerformance WHERE `Average Rating` BETWEEN 3.5 AND 4;

```
hive> SELECT DISTINCT `Agent Name` FROM AgentPerformance WHERE `Average Rating` BETWEEN 3.5 AND 4;
Query ID = cloudera_20220930035151_e74debac-0f2f-43cf-afbd-1a7495dfc0f9
Total jobs = 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_1664521724151_0020, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1664521724151_0020/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1664521724151_0020
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-09-30 03:51:57,944 Stage-1 map = 0%,  reduce = 0%
2022-09-30 03:52:04,184 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.89 sec
2022-09-30 03:52:10,675 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 3.42 sec
MapReduce Total cumulative CPU time: 3 seconds 420 msec
Ended Job = job_1664521724151_0020
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 3.42 sec   HDFS Read: 119429 HDFS Write: 541 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 420 msec
OK
agent name
Aditya Shinde
Aditya_iot
Ameya Jain
Anirudh
Anurag Tiwari
Ashad Nasim
Ayushi Mishra
Boktiar Ahmed Bappy
Chaitra K Hiremath
Deepranjan Gupta
Harikrishnan Shaji
Hrisikesh Neogi
Ishawant Kumar
Jawala Prakash
Jayant Kumar
Jaydeep Dixit
Khushboo Priya
Madhulika G
Mahesh Sarade
Maitry
Maneesh
Manjunatha A
Mithun S
Muskan Garg
Nandani Gupta
Nishtha Jain
```

```
2022-09-30 03:52:10,675 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 3.42 sec
MapReduce Total cumulative CPU time: 3 seconds 420 msec
Ended Job = job_1664521724151_0020
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 3.42 sec   HDFS Read: 119429 HDFS Write: 541 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 420 msec
OK
agent name
Aditya Shinde
Aditya_iot
Ameya Jain
Anirudh
Anurag Tiwari
Ashad Nasim
Ayushi Mishra
Boktiar Ahmed Bappy
Chaitra K Hiremath
Deepranjan Gupta
Harikrishnan Shaji
Hrisikesh Neogi
Ishawant Kumar
Jawala Prakash
Jayant Kumar
Jaydeep Dixit
Khushboo Priya
Madhulika G
Mahesh Sarade
Maitry
Maneesh
Manjunatha A
Mithun S
Muskan Garg
Nandani Gupta
Nishtha Jain
Prabir Kumar Satapathy
Prateek _iot
Prerna Singh
Rishav Dash
Sandipan Saha
Sanjeev Kumar
Saurabh Shukla
Shiva Srivastava
Shivan K
Shivananda Sonwane
Shubham Sharma
Swati
Wasim
Zeeshan
Time taken: 22.863 seconds, Fetched: 40 row(s)
hive>
```

9. Agent name who have rating less than 3.5.

SELECT DISTINCT `Agent Name` FROM AgentPerformance WHERE `Average Rating` < 3.5;

```
hive> SELECT DISTINCT `Agent Name` FROM AgentPerformance WHERE `Average Rating` < 3.5;
Query ID = cloudera_20220930034848_fbb3cbb1-7456-4008-977a-c6eb70f2c64a
Total jobs = 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_1664521724151_0019, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1664521724151_0019/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1664521724151_0019
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-09-30 03:48:24,275 Stage-1 map = 0%,  reduce = 0%
2022-09-30 03:48:32,684 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.98 sec
2022-09-30 03:48:40,525 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 4.18 sec
MapReduce Total cumulative CPU time: 4 seconds 180 msec
Ended Job = job_1664521724151_0019
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 4.18 sec   HDFS Read: 119003 HDFS Write: 867 SUCCESS
Total MapReduce CPU Time Spent: 4 seconds 180 msec
OK
agent name
Abhishek
Aditya
Aditya Shinde
Aditya_iot
Amersh
Ameya Jain
Anirudh
Ankit Sharma
Ankitjha
Anurag Tiwari
Aravind
Ashad Nasim
Ashish
Ayushi Mishra
Bharath
Boktiar Ahmed Bappy
Chaitra K Hiremath
Deepranjan Gupta
Dibyanshu
Harikrishnan Shaji
Hitesh Choudhary
Hrisikesh Neogi
Hyder Abbas
Ineuron Intelligence
Ishawant Kumar
Jawala Prakash
```

```
Hyder Abbas
Ineuron Intelligence
Ishawant Kumar
Jawala Prakash
Jayant Kumar
Jaydeep Dixit
Khushboo Priya
Madhulika G
Mahak
Mahesh Sarade
Maitry
Maneesh
Manjunatha A
Mithun S
Mukesh
Mukesh Rao
Muskan Garg
Nandani Gupta
Nishtha Jain
Nitin M
Prabir Kumar Satapathy
Prateek _iot
Prerna Singh
Rishav Dash
Rohan
Saif Khan
Saikumarreddy N
Samprit
Sandipan Saha
Sanjeev Kumar
Sanjeevan
Saurabh Shukla
Shiva Srivastava
Shivan K
Shivan_S
Shivananda Sonwane
Shubham Sharma
Sowmiya Sivakumar
Spuri
Sudhanshu Kumar
Suraj S Bilgi
Swati
Tarun
Uday Mishra
Vasanth P
Vivek
Wasim
Zeeshan
Time taken: 25.38 seconds, Fetched: 70 row(s)
hive>
```

10. Agent name who have rating more than 4.5

SELECT DISTINCT `Agent Name` FROM AgentPerformance WHERE `Average Rating` > 4.5;

```
hive> SELECT DISTINCT `Agent Name` FROM AgentPerformance WHERE `Average Rating` > 4.5;
Query ID = cloudera_20220930035353_955d0f67-bdf4-4583-a891-b23855eefbd8
Total jobs = 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_1664521724151_0021, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1664521724151_0021/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1664521724151_0021
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-09-30 03:53:54,662 Stage-1 map = 0%,  reduce = 0%
2022-09-30 03:54:01,081 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.84 sec
2022-09-30 03:54:07,466 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 3.58 sec
MapReduce Total cumulative CPU time: 3 seconds 580 msec
Ended Job = job_1664521724151_0021
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 3.58 sec   HDFS Read: 119015 HDFS Write: 625 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 580 msec
OK
agent name
Aditya Shinde
Aditya_iot
Ameya Jain
Anirudh
Ankitjha
Aravind
Ayushi Mishra
Bharath
Boktiar Ahmed Bappy
Chaitra K Hiremath
Deepranjan Gupta
Harikrishnan Shaji
Hrisikesh Neogi
Ishawant Kumar
Jawala Prakash
Jayant Kumar
Jaydeep Dixit
Khushboo Priya
Madhulika G
Mahesh Sarade
Maitry
Manjunatha A
Mithun S
Mukesh
Mukesh Rao
Muskan Garg
```

```
OK
Aditya Shinde
Aditya_iot
Ameya Jain
Anirudh
Ankitjha
Aravind
Ayushi Mishra
Bharath
Boktiar Ahmed Bappy
Chaitra K Hiremath
Deepranjan Gupta
Harikrishnan Shaji
Hrisikesh Neogi
Ishawant Kumar
Jawala Prakash
Jayant Kumar
Jaydeep Dixit
Khushboo Priya
Madhulika G
Mahesh Sarade
Maitry
Manjunatha A
Mithun S
Mukesh
Mukesh Rao
Muskan Garg
Nandani Gupta
Nishtha Jain
Prabir Kumar Satapathy
Prateek _iot
Prerna Singh
Rishav Dash
Saikumarreddy N
Sandipan Saha
Sanjeev Kumar
Saurabh Shukla
Shiva Srivastava
Shivan K
Shivananda Sonwane
Shubham Sharma
Sowmiya Sivakumar
Sudhanshu Kumar
Suraj S Bilgi
Swati
Vivek
Wasim
Zeeshan
Time taken: 23.36 seconds, Fetched: 47 row(s)
hive>
```

11. How many feedback agents have received more than 4.5 average.

SELECT SUM(`Total Feedback`) AS `Total Feedback` FROM AgentPerformance WHERE `Average Rating` > 4.5;

```
hive> SELECT SUM(`Total Feedback`) AS `Total Feedback` FROM AgentPerformance WHERE `Average Rating` > 4.5;
Query ID = cloudera_20220930035656_82e525ac-6c39-43ee-a2b9-9639cf546c6c
Total jobs = 1
Launching Job 1 out of 1
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>
Starting Job = job_1664521724151_0022, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1664521724151_0022/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1664521724151_0022
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-09-30 03:56:39,602 Stage-1 map = 0%,  reduce = 0%
2022-09-30 03:56:45,849 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.71 sec
2022-09-30 03:56:53,161 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 3.57 sec
MapReduce Total cumulative CPU time: 3 seconds 570 msec
Ended Job = job_1664521724151_0022
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 3.57 sec   HDFS Read: 119354 HDFS Write: 5 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 570 msec
OK
total feedback
3489
Time taken: 20.792 seconds, Fetched: 1 row(s)
hive>
```

12. average weekly response time for each agent.

hive> select `agent name`, avg(time[0]*60*60+time[1]*60+time[2]) as avg_weekly_response_time_in_seconds from
> (select `agent name`, split(`average response time`,':') as time from agentperformance) s group by `agent name`;

```
hive> select `agent name`, avg(time[0]*60*60+time[1]*60+time[2]) as avg_weekly_response_time_in_seconds from
    > (select `agent name`, split(`average response time`,':') as time from agentperformance) s group by `agent name`;
Query ID = cloudera_20221011033737_9fce5484-fc2c-4a00-b5fd-e673cebbf8d5
Total jobs = 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_1665480075118_0007, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1665480075118_0007/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1665480075118_0007
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-11 03:37:22,599 Stage-1 map = 0%,  reduce = 0%
2022-10-11 03:37:35,053 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 3.96 sec
2022-10-11 03:37:47,685 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 7.23 sec
MapReduce Total cumulative CPU time: 7 seconds 230 msec
Ended Job = job_1665480075118_0007
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 7.23 sec   HDFS Read: 122132 HDFS Write: 1643 SUCCESS
Total MapReduce CPU Time Spent: 7 seconds 230 msec
OK
agent name      avg_weekly_response_time_in_seconds
Abhishek        0.0
Aditya  0.0
Aditya Shinde   29.733333333333334
Aditya_iot      33.96666666666667
Amersh  0.0
Ameya Jain      21.133333333333333
Anirudh         21.766666666666666
Ankit Sharma    0.0
Ankitjha        4.433333333333334
Anurag Tiwari   8.433333333333334
Aravind         21.366666666666667
Ashad Nasim     38.63333333333333
Ashish  0.0
Ayushi Mishra   60.333333333333336
Bharath         26.8
Boktiar Ahmed Bappy     66.03333333333333
Chaitra K Hiremath      15.133333333333333
Deepranjan Gupta        53.2
Dibyanshu       1.2666666666666666
Harikrishnan Shaji      33.96666666666667
Hitesh Choudhary        0.0
Hrisikesh Neogi 50.5
Hyder Abbas     0.0
Ineuron Intelligence    0.0
Ishawant Kumar  50.13333333333333
```

13. average weekly resolution time for each agents

hive> select `agent name`, avg(time[0]*60*60+time[1]*60+time[2]) as avg_weekly_resolution_time_in_seconds from
> (select `agent name`, split(`average resolution time`,':') as time from agentperformance) s group by `agent name`;

```
hive> select `agent name`, avg(time[0]*60*60+time[1]*60+time[2]) as avg_weekly_resolution_time_in_seconds from
    > (select `agent name`, split(`average resolution time`,':') as time from agentperformance) s group by `agent name`;
Query ID = cloudera_20221011034040_670ae047-e49f-4822-b0e7-e4daf638a0d7
Total jobs = 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_1665480075118_0008, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1665480075118_0008/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1665480075118_0008
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-11 03:41:08,908 Stage-1 map = 0%,  reduce = 0%
2022-10-11 03:41:17,760 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 3.05 sec
2022-10-11 03:41:27,853 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 5.97 sec
MapReduce Total cumulative CPU time: 5 seconds 970 msec
Ended Job = job_1665480075118_0008
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 5.97 sec   HDFS Read: 122125 HDFS Write: 1760 SUCCESS
Total MapReduce CPU Time Spent: 5 seconds 970 msec
OK
agent name      avg_weekly_resolution_time_in_seconds
Abhishek        0.0
Aditya  0.0
Aditya Shinde   620.6333333333333
Aditya_iot      589.3
Amersh  0.0
Ameya Jain      328.5
Anirudh         185.46666666666667
Ankit Sharma    0.0
Ankitjha        54.46666666666667
Anurag Tiwari   73.93333333333334
Aravind         535.5666666666667
Ashad Nasim     20.933333333333334
Ashish  0.0
Ayushi Mishra   922.6
Bharath         647.6
Boktiar Ahmed Bappy     1023.8333333333334
Chaitra K Hiremath      88.96666666666667
Deepranjan Gupta        1251.1666666666667
Dibyanshu       24.666666666666668
Harikrishnan Shaji      678.4333333333333
Hitesh Choudhary        2.8333333333333335
Hrisikesh Neogi 926.3
Hyder Abbas     0.0
Ineuron Intelligence    0.0
Ishawant Kumar  865.1
```

14. Find the number of chat on which they have received a feedback
SELECT SUM(`Total chats`) AS Chats FROM AgentPerformance WHERE `Total feedback` != 0;

```
hive> SELECT SUM(`Total chats`) AS Chats FROM AgentPerformance WHERE `Total feedback` != 0;
Query ID = cloudera_20220930040101_095b4fee-d30b-41b5-984a-7813a6301c70
Total jobs = 1
Launching Job 1 out of 1
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>
Starting Job = job_1664521724151_0023, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1664521724151_0023/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1664521724151_0023
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-09-30 04:01:25,693 Stage-1 map = 0%,  reduce = 0%
2022-09-30 04:01:32,077 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 2.1 sec
2022-09-30 04:01:39,435 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 3.87 sec
MapReduce Total cumulative CPU time: 3 seconds 870 msec
Ended Job = job_1664521724151_0023
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 3.87 sec   HDFS Read: 119335 HDFS Write: 6 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 870 msec
OK
chats
14702
Time taken: 22.025 seconds, Fetched: 1 row(s)
hive>
```

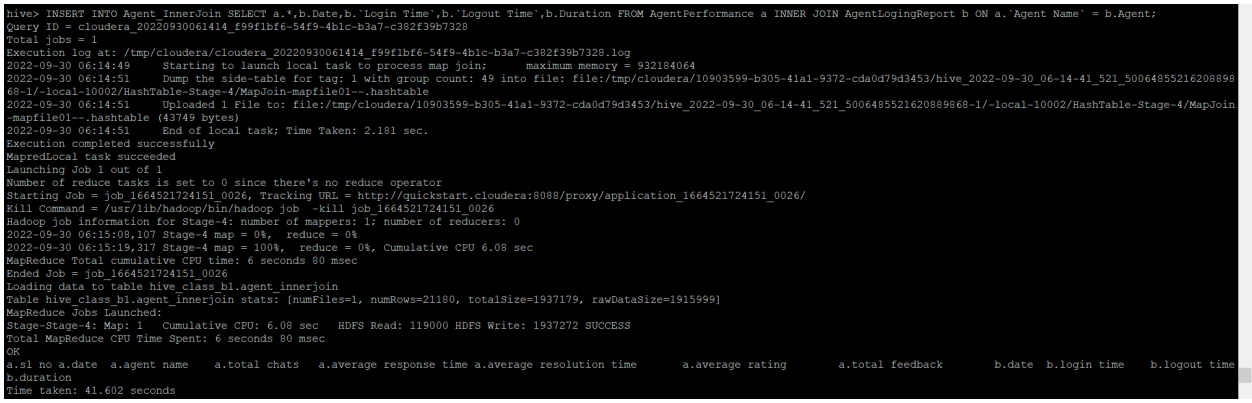
15. Total contribution hour for each and every agents weekly basis

16. Perform inner join, left join and right join based on the agent column and after joining the table export that data into your local system.

INNER JOIN

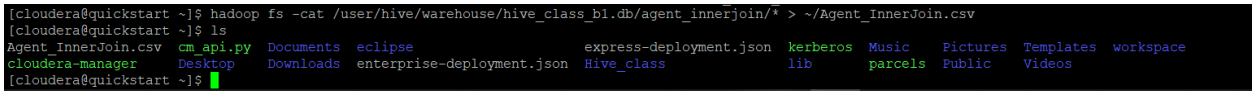
```
hive> CREATE TABLE IF NOT EXISTS Agent_InnerJoin
    > (
    > `SL No` INT,
    > `a_Date` STRING,
    > `Agent Name` STRING,
    > `Total Chats` INT,
    > `Average Response Time` STRING,
    > `Average Resolution Time` STRING,
    > `Average Rating` FLOAT,
    > `Total Feedback` INT,
    > `b_Date` STRING,
    > `Login Time` STRING,
    > `Logout Time` STRING,
    > Duration STRING
    > )
    > row format delimited
    > fields terminated by ','
    > stored as textfile;
OK
Time taken: 0.485 seconds
```

INSERT INTO Agent_InnerJoin SELECT a.*,b.Date,b.`Login Time`,b.`Logout Time`,b.Duration FROM AgentPerformance a INNER JOIN AgentLogingReport b ON a.`Agent Name` = b.Agent;



```
hive> INSERT INTO Agent_InnerJoin SELECT a.*,b.Date,b.`Login Time`,b.`Logout Time`,b.Duration FROM AgentPerformance a INNER JOIN AgentLogingReport b ON a.`Agent Name` = b.Agent;
Query ID = cloudera_20220930061414_f99f1bf6-54f9-4b1c-b3a7-c382f39b7328
Total jobs = 1
Execution log at: /tmp/cloudera/cloudera_20220930061414_f99f1bf6-54f9-4b1c-b3a7-c382f39b7328.log
2022-09-30 06:14:49     Starting to launch local task to process map join;      maximum memory = 932184064
2022-09-30 06:14:51     Dump the side-table for tag: 1 with group count: 49 into file: file:/tmp/cloudera/10903599-b305-41a1-9372-cda0d79d3453/hive_2022-09-30_06-14-41_521_5006485521620889868-1/-local-10002/HashTable-Stage-4/MapJoin-mapfile01--.hashtable
2022-09-30 06:14:51     Uploaded 1 File to: file:/tmp/cloudera/10903599-b305-41a1-9372-cda0d79d3453/hive_2022-09-30_06-14-41_521_5006485521620889868-1/-local-10002/HashTable-Stage-4/MapJoin-mapfile01--.hashtable (43749 bytes)
2022-09-30 06:14:51     End of local task; Time Taken: 2.181 sec.
Execution completed successfully
MapredLocal task succeeded
Launching Job 1 out of 1
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1664521724151_0026, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1664521724151_0026/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1664521724151_0026
Hadoop job information for Stage-4: number of mappers: 1; number of reducers: 0
2022-09-30 06:15:08,107 Stage-4 map = 0%,  reduce = 0%
2022-09-30 06:15:19,317 Stage-4 map = 100%,  reduce = 0%, Cumulative CPU 6.08 sec
MapReduce Total cumulative CPU time: 6 seconds 80 msec
Ended Job = job_1664521724151_0026
Loading data to table hive_class_b1.agent_innerjoin
Table hive_class_b1.agent_innerjoin stats: [numFiles=1, numRows=21180, totalSize=1937179, rawDataSize=1915999]
MapReduce Jobs Launched:
Stage-Stage-4: Map: 1   Cumulative CPU: 6.08 sec   HDFS Read: 119000 HDFS Write: 1937272 SUCCESS
Total MapReduce CPU Time Spent: 6 seconds 80 msec
OK
a.sl no a.date  a.agent name    a.total chats   a.average response time a.average resolution time       a.average rating        a.total feedback        b.date  b.login time    b.logout time   b.duration
Time taken: 41.602 seconds
```

hadoop fs -cat /user/hive/warehouse/hive_class_b1.db/agent_innerjoin/* > ~/Agent_InnerJoin.csv



```
[cloudera@quickstart ~]$ hadoop fs -cat /user/hive/warehouse/hive_class_b1.db/agent_innerjoin/* > ~/Agent_InnerJoin.csv
[cloudera@quickstart ~]$ ls
Agent_InnerJoin.csv  cm_api.py  Documents  eclipse                  express-deployment.json  kerberos  Music    Pictures  Templates  workspace
cloudera-manager     Desktop    Downloads  enterprise-deployment.json  Hive_class               lib       parcels  Public    Videos
[cloudera@quickstart ~]$
```

LEFT JOIN

```
hive> CREATE TABLE IF NOT EXISTS Agent_LeftJoin
    > (
    > `SL No` INT,
    > `a_Date` STRING,
    > `Agent Name` STRING,
    > `Total Chats` INT,
    > `Average Response Time` STRING,
    > `Average Resolution Time` STRING,
    > `Average Rating` FLOAT,
    > `Total Feedback` INT,
    > `b_Date` STRING,
    > `Login Time` STRING,
    > `Logout Time` STRING,
    > Duration STRING
    > )
    > row format delimited
    > fields terminated by ','
    > stored as textfile;
OK
Time taken: 0.842 seconds
hive>
```

INSERT INTO Agent_LeftJoin SELECT a.*,b.Date,b.`Login Time`,b.`Logout Time`,b.Duration FROM AgentPerformance a LEFT JOIN AgentLogingReport b ON a.`Agent Name` = b.Agent;

```
hive> INSERT INTO Agent_LeftJoin SELECT a.*,b.Date,b.`Login Time`,b.`Logout Time`,b.Duration FROM AgentPerformance a LEFT JOIN AgentLogingReport b ON a.`Agent Name` = b.Agent;
Query ID = cloudera_20220930063838_f94b7b1a-7266-446e-b49e-330e7969de8f
Total jobs = 1
Execution log at: /tmp/cloudera/cloudera_20220930063838_f94b7b1a-7266-446e-b49e-330e7969de8f.log
2022-09-30 06:38:42     Starting to launch local task to process map join;      maximum memory = 932184064
2022-09-30 06:38:44     Dump the side-table for tag: 1 with group count: 49 into file: file:/tmp/cloudera/272e5e46-0405-4bd0-b640-fc7b3aca2f47/hive_2022-09-30_06-38-34_547_192766224510318689
99-1/-local-10002/HashTable-Stage-4/MapJoin-mapfile01--.hashtable
2022-09-30 06:38:44     Uploaded 1 File to: file:/tmp/cloudera/272e5e46-0405-4bd0-b640-fc7b3aca2f47/hive_2022-09-30_06-38-34_547_1927662245103186899-1/-local-10002/HashTable-Stage-4/MapJoin
-mapfile01--.hashtable (43749 bytes)
2022-09-30 06:38:44     End of local task; Time Taken: 1.809 sec.
Execution completed successfully
MapredLocal task succeeded
Launching Job 1 out of 1
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1664521724151_0027, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1664521724151_0027/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1664521724151_0027
Hadoop job information for Stage-4: number of mappers: 1; number of reducers: 0
2022-09-30 06:39:01,692 Stage-4 map = 0%,  reduce = 0%
2022-09-30 06:39:12,009 Stage-4 map = 100%,  reduce = 0%, Cumulative CPU 4.48 sec
MapReduce Total cumulative CPU time: 4 seconds 480 msec
Ended Job = job_1664521724151_0027
Loading data to table hive_class_b1.agent_leftjoin
Table hive_class_b1.agent_leftjoin stats: [numFiles=1, numRows=22260, totalSize=2002671, rawDataSize=1980411]
MapReduce Jobs Launched:
Stage-Stage-4: Map: 1   Cumulative CPU: 4.48 sec   HDFS Read: 118787 HDFS Write: 2002763 SUCCESS
Total MapReduce CPU Time Spent: 4 seconds 480 msec
OK
Time taken: 40.9 seconds
hive>
```

hadoop fs -cat /user/hive/warehouse/hive_class_b1.db/agent_leftjoin/* > ~/Agent_LeftJoin.csv

```
[cloudera@quickstart ~]$ hadoop fs -cat /user/hive/warehouse/hive_class_b1.db/agent_leftjoin/* > ~/Agent_LeftJoin.csv
[cloudera@quickstart ~]$ ls
Agent_InnerJoin.csv  cloudera-manager  Desktop    Downloads  enterprise-deployment.json  Hive_class  lib    parcels   Public     Videos
Agent_LeftJoin.csv   cm_api.py         Documents  eclipse    express-deployment.json     kerberos    Music  Pictures  Templates  workspace
[cloudera@quickstart ~]$
```

RIGHT JOIN

```
hive> CREATE TABLE IF NOT EXISTS Agent_RightJoin
    > (
    > `SL No` INT,
    > `a_Date` STRING,
    > `Agent Name` STRING,
    > `Total Chats` INT,
    > `Average Response Time` STRING,
    > `Average Resolution Time` STRING,
    > `Average Rating` FLOAT,
    > `Total Feedback` INT,
    > `b_Date` STRING,
    > `Login Time` STRING,
    > `Logout Time` STRING,
    > Duration STRING
    > )
    > row format delimited
    > fields terminated by ','
    > stored as textfile;
OK
Time taken: 0.799 seconds
hive>
```

INSERT INTO Agent_RightJoin SELECT a.*,b.Date,b.`Login Time`,b.`Logout Time`,b.Duration FROM AgentPerformance a RIGHT JOIN AgentLogingReport b ON a.`Agent Name` = b.Agent;

```
hive> INSERT INTO Agent_RightJoin SELECT a.*,b.Date,b.`Login Time`,b.`Logout Time`,b.Duration FROM AgentPerformance a RIGHT JOIN AgentLogingReport b ON a.`Agent Name` = b.Agent;
Query ID = cloudera_20220930064747_0ff0ee6a-6c5f-40ce-a877-889a075929e3
Total jobs = 1
Execution log at: /tmp/cloudera/cloudera_20220930064747_0ff0ee6a-6c5f-40ce-a877-889a075929e3.log
2022-09-30 06:47:11     Starting to launch local task to process map join;      maximum memory = 932184064
2022-09-30 06:47:13     Dump the side-table for tag: 0 with group count: 70 into file: file:/tmp/cloudera/c4ff2681-1623-4389-b5eb-c614ea3295f5/hive_2022-09-30_06-47-03_683_3225855369152444359-1/-local-10002/HashTable-Stage-4/MapJoin-mapfile00--.hashtable
2022-09-30 06:47:13     Uploaded 1 File to: file:/tmp/cloudera/c4ff2681-1623-4389-b5eb-c614ea3295f5/hive_2022-09-30_06-47-03_683_3225855369152444359-1/-local-10002/HashTable-Stage-4/MapJoin-mapfile00--.hashtable (87832 bytes)
2022-09-30 06:47:13     End of local task; Time Taken: 2.584 sec.
Execution completed successfully
MapredLocal task succeeded
Launching Job 1 out of 1
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1664521724151_0028, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1664521724151_0028/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1664521724151_0028
Hadoop job information for Stage-4: number of mappers: 1; number of reducers: 0
2022-09-30 06:47:28,588 Stage-4 map = 0%,  reduce = 0%
2022-09-30 06:47:40,077 Stage-4 map = 100%,  reduce = 0%, Cumulative CPU 4.83 sec
MapReduce Total cumulative CPU time: 4 seconds 830 msec
Ended Job = job_1664521724151_0028
Loading data to table hive_class_b1.agent_rightjoin
Table hive_class_b1.agent_rightjoin stats: [numFiles=1, numRows=21486, totalSize=1955846, rawDataSize=1934360]
MapReduce Jobs Launched:
Stage-Stage-4: Map: 1   Cumulative CPU: 4.83 sec   HDFS Read: 64322 HDFS Write: 1955939 SUCCESS
Total MapReduce CPU Time Spent: 4 seconds 830 msec
OK
a.sl no a.date  a.agent name    a.total chats   a.average response time a.average resolution time       a.average rating        a.total feedback        b.date  b.login time    b.logout time   b.duration
Time taken: 39.936 seconds
hive>
```

hadoop fs -cat /user/hive/warehouse/hive_class_b1.db/agent_rightjoin/* > ~/Agent_RightJoin.csv

```
[cloudera@quickstart ~]$ hadoop fs -cat /user/hive/warehouse/hive_class_b1.db/agent_rightjoin/* > ~/Agent_RightJoin.csv
[cloudera@quickstart ~]$ ls
Agent_InnerJoin.csv  Agent_RightJoin.csv  cm_api.py  Documents  eclipse                  express-deployment.json  kerberos  Music    Pictures  Templates  workspace
Agent_LeftJoin.csv   cloudera-manager     Desktop    Downloads  enterprise-deployment.json  Hive_class               lib       parcels  Public    Videos
[cloudera@quickstart ~]$
```

17. Perform partitioning on top of the agent column and then on top of that perform bucketing for each partitioning.

```
hive> set hive.exec.dynamic.partition.mode = nonstrict;
hive> set hive.enforce.bucketing = true;
hive> CREATE TABLE IF NOT EXISTS AgentPerformance_Part_Buck
    > (
    > `SL No` INT,
    > `Date` STRING,
    > `Total Chats` INT,
    > `Average Response Time` STRING,
    > `Average Resolution Time` STRING,
    > `Average Rating` FLOAT,
    > `Total Feedback` INT
    > )
    > partitioned by(`Agent Name` STRING)
    > clustered by(`SL No`)
    > into 4 buckets;
OK
Time taken: 0.486 seconds
hive> INSERT OVERWRITE TABLE AgentPerformance_Part_Buck Partition(`Agent Name`)
    > SELECT `SL No`,
    > `Date`,
    > `Total Chats`,
    > `Average Response Time`,
    > `Average Resolution Time`,
    > `Average Rating`,
    > `Total Feedback`,
    > `Agent Name`
    > FROM AgentPerformance;
Query ID = cloudera_20220930090000_45b1aa06-e8d1-4d91-a081-af90b3b27a20
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 4
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_1664521724151_0036, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1664521724151_0036/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1664521724151_0036
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 4
2022-09-30 09:00:51,521 Stage-1 map = 0%,  reduce = 0%
2022-09-30 09:01:00,765 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 3.12 sec
2022-09-30 09:01:41,771 Stage-1 map = 100%,  reduce = 17%, Cumulative CPU 6.54 sec
2022-09-30 09:01:51,694 Stage-1 map = 100%,  reduce = 50%, Cumulative CPU 12.49 sec
2022-09-30 09:02:27,306 Stage-1 map = 100%,  reduce = 67%, Cumulative CPU 29.42 sec
2022-09-30 09:02:30,575 Stage-1 map = 100%,  reduce = 75%, Cumulative CPU 42.26 sec
2022-09-30 09:02:37,630 Stage-1 map = 100%,  reduce = 92%, Cumulative CPU 50.19 sec
2022-09-30 09:02:50,667 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 62.31 sec
MapReduce Total cumulative CPU time: 1 minutes 12 seconds 300 msec
```

```
Partition hive_class_b1.agentperformance_part_buck{agent name=Jayant Kumar} stats: [numFiles=4, numRows=30, totalSize=1160, rawDataSize=1130]
Partition hive_class_b1.agentperformance_part_buck{agent name=Jaydeep Dixit} stats: [numFiles=4, numRows=30, totalSize=1203, rawDataSize=1173]
Partition hive_class_b1.agentperformance_part_buck{agent name=Khushboo Priya} stats: [numFiles=4, numRows=30, totalSize=1200, rawDataSize=1170]
Partition hive_class_b1.agentperformance_part_buck{agent name=Madhulika G} stats: [numFiles=4, numRows=30, totalSize=1201, rawDataSize=1171]
Partition hive_class_b1.agentperformance_part_buck{agent name=Mahak } stats: [numFiles=4, numRows=30, totalSize=1146, rawDataSize=1116]
Partition hive_class_b1.agentperformance_part_buck{agent name=Mahesh Sarade} stats: [numFiles=4, numRows=30, totalSize=1184, rawDataSize=1154]
Partition hive_class_b1.agentperformance_part_buck{agent name=Maitry } stats: [numFiles=4, numRows=30, totalSize=1198, rawDataSize=1168]
Partition hive_class_b1.agentperformance_part_buck{agent name=Maneesh } stats: [numFiles=4, numRows=30, totalSize=1146, rawDataSize=1116]
Partition hive_class_b1.agentperformance_part_buck{agent name=Manjunatha A} stats: [numFiles=4, numRows=30, totalSize=1198, rawDataSize=1168]
Partition hive_class_b1.agentperformance_part_buck{agent name=Mithun S} stats: [numFiles=4, numRows=30, totalSize=1188, rawDataSize=1158]
Partition hive_class_b1.agentperformance_part_buck{agent name=Mukesh } stats: [numFiles=4, numRows=30, totalSize=1147, rawDataSize=1117]
Partition hive_class_b1.agentperformance_part_buck{agent name=Mukesh Rao } stats: [numFiles=4, numRows=30, totalSize=1147, rawDataSize=1117]
Partition hive_class_b1.agentperformance_part_buck{agent name=Muskan Garg} stats: [numFiles=4, numRows=30, totalSize=1152, rawDataSize=1122]
Partition hive_class_b1.agentperformance_part_buck{agent name=Nandani Gupta} stats: [numFiles=4, numRows=30, totalSize=1202, rawDataSize=1172]
Partition hive_class_b1.agentperformance_part_buck{agent name=Nishtha Jain} stats: [numFiles=4, numRows=30, totalSize=1188, rawDataSize=1158]
Partition hive_class_b1.agentperformance_part_buck{agent name=Nitin M} stats: [numFiles=4, numRows=30, totalSize=1146, rawDataSize=1116]
Partition hive_class_b1.agentperformance_part_buck{agent name=Prabir Kumar Satapathy} stats: [numFiles=4, numRows=30, totalSize=1179, rawDataSize=1149]
Partition hive_class_b1.agentperformance_part_buck{agent name=Prateek __iot } stats: [numFiles=4, numRows=30, totalSize=1168, rawDataSize=1138]
Partition hive_class_b1.agentperformance_part_buck{agent name=Prerna Singh} stats: [numFiles=4, numRows=30, totalSize=1189, rawDataSize=1159]
Partition hive_class_b1.agentperformance_part_buck{agent name=Rishav Dash} stats: [numFiles=4, numRows=60, totalSize=2341, rawDataSize=2281]
Partition hive_class_b1.agentperformance_part_buck{agent name=Rohan } stats: [numFiles=4, numRows=30, totalSize=1146, rawDataSize=1116]
Partition hive_class_b1.agentperformance_part_buck{agent name=Saif Khan} stats: [numFiles=4, numRows=30, totalSize=1146, rawDataSize=1116]
Partition hive_class_b1.agentperformance_part_buck{agent name=Saikumarreddy N} stats: [numFiles=4, numRows=30, totalSize=1182, rawDataSize=1152]
Partition hive_class_b1.agentperformance_part_buck{agent name=Samprit } stats: [numFiles=4, numRows=30, totalSize=1146, rawDataSize=1116]
Partition hive_class_b1.agentperformance_part_buck{agent name=Sandipan Saha} stats: [numFiles=4, numRows=30, totalSize=1150, rawDataSize=1120]
Partition hive_class_b1.agentperformance_part_buck{agent name=Sanjeev Kumar} stats: [numFiles=4, numRows=30, totalSize=1198, rawDataSize=1168]
Partition hive_class_b1.agentperformance_part_buck{agent name=Sanjeevan } stats: [numFiles=4, numRows=30, totalSize=1145, rawDataSize=1115]
Partition hive_class_b1.agentperformance_part_buck{agent name=Saurabh Shukla} stats: [numFiles=4, numRows=30, totalSize=1147, rawDataSize=1117]
Partition hive_class_b1.agentperformance_part_buck{agent name=Shiva Srivastava} stats: [numFiles=4, numRows=30, totalSize=1150, rawDataSize=1120]
Partition hive_class_b1.agentperformance_part_buck{agent name=Shivan K} stats: [numFiles=4, numRows=30, totalSize=1187, rawDataSize=1157]
Partition hive_class_b1.agentperformance_part_buck{agent name=Shivan_S } stats: [numFiles=4, numRows=30, totalSize=1147, rawDataSize=1117]
Partition hive_class_b1.agentperformance_part_buck{agent name=Shivananda Sonwane} stats: [numFiles=4, numRows=30, totalSize=1198, rawDataSize=1168]
Partition hive_class_b1.agentperformance_part_buck{agent name=Shubham Sharma} stats: [numFiles=4, numRows=30, totalSize=1199, rawDataSize=1169]
Partition hive_class_b1.agentperformance_part_buck{agent name=Sowmiya Sivakumar} stats: [numFiles=4, numRows=30, totalSize=1171, rawDataSize=1141]
Partition hive_class_b1.agentperformance_part_buck{agent name=Spuri } stats: [numFiles=4, numRows=30, totalSize=1146, rawDataSize=1116]
Partition hive_class_b1.agentperformance_part_buck{agent name=Sudhanshu Kumar} stats: [numFiles=4, numRows=30, totalSize=1147, rawDataSize=1117]
Partition hive_class_b1.agentperformance_part_buck{agent name=Suraj S Bilgi} stats: [numFiles=4, numRows=30, totalSize=1148, rawDataSize=1118]
Partition hive_class_b1.agentperformance_part_buck{agent name=Swati } stats: [numFiles=4, numRows=30, totalSize=1191, rawDataSize=1161]
Partition hive_class_b1.agentperformance_part_buck{agent name=Tarun } stats: [numFiles=4, numRows=30, totalSize=1147, rawDataSize=1117]
Partition hive_class_b1.agentperformance_part_buck{agent name=Uday Mishra} stats: [numFiles=4, numRows=30, totalSize=1146, rawDataSize=1116]
Partition hive_class_b1.agentperformance_part_buck{agent name=Vasanth P} stats: [numFiles=4, numRows=30, totalSize=1146, rawDataSize=1116]
Partition hive_class_b1.agentperformance_part_buck{agent name=Vivek } stats: [numFiles=4, numRows=30, totalSize=1150, rawDataSize=1120]
Partition hive_class_b1.agentperformance_part_buck{agent name=Wasim } stats: [numFiles=4, numRows=30, totalSize=1190, rawDataSize=1160]
Partition hive_class_b1.agentperformance_part_buck{agent name=Zeeshan } stats: [numFiles=4, numRows=30, totalSize=1192, rawDataSize=1162]
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 4   Cumulative CPU: 72.3 sec   HDFS Read: 135108 HDFS Write: 107890 SUCCESS
Total MapReduce CPU Time Spent: 1 minutes 12 seconds 300 msec
OK
Time taken: 199.426 seconds
hive>
```