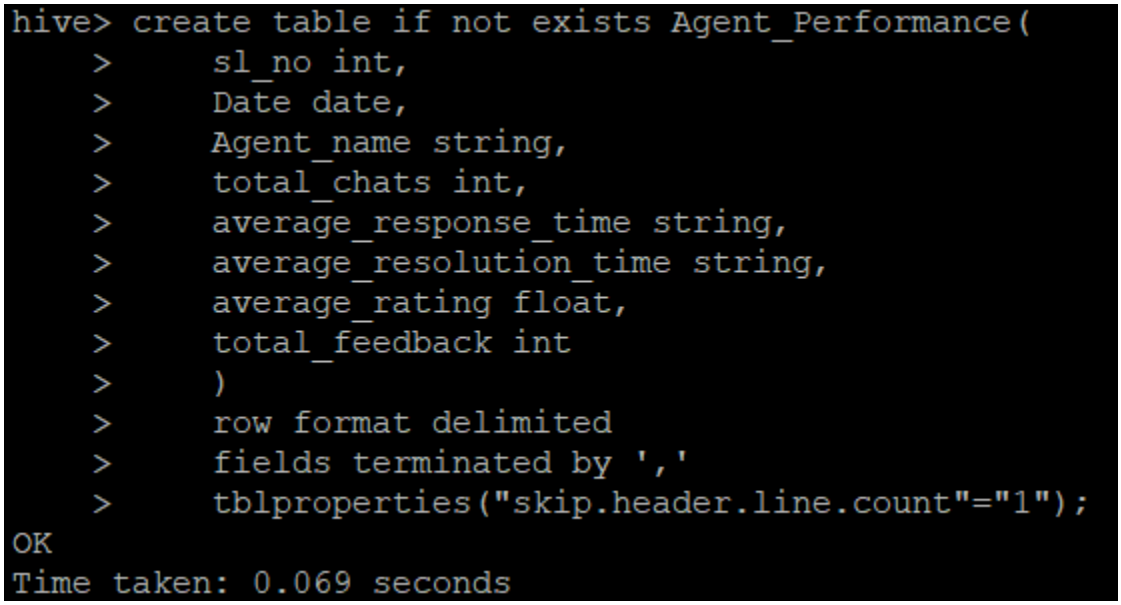
Hive Mini Project-1:--

==========

1. Create a schema based on the given dataset.

```
hive> create table if not exists Agent_Performance(
   >    sl_no int,
   >    Date date,
   >    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: 0.069 seconds
```




```
hive> create table if not exists AgentLogingReport(
   > sl_no int,
   > Agent_name string,
   > Date date,
   > login_time string,
   > logout_time string,
   > duration string)
   > row format delimited
```

```
   > fields terminated by ','
   > tblproperties("skip.header.line.count"="1");
OK
Time taken: 0.081 seconds
```

```
hive> create table if not exists AgentLogingReport(
    > sl_no int,
    > Agent_name string,
    > Date date,
    > login_time string,
    > logout_time string,
    > duration string)
    > row format delimited
    > fields terminated by ','
    > tblproperties("skip.header.line.count"="1");
OK
Time taken: 0.081 seconds
```

2.Dump the data in side the HDFS in the given schema location

hive> load data local inpath 'file:///tmp/hive_class_1/AgentPerformance.csv' into table Agent_Performance;
Loading data to table mini_project_1.agent_performance
Table mini_project_1.agent_performance stats: [numFiles=1, totalSize=112661]
OK
Time taken: 0.516 seconds

```
hive> load data local inpath 'file:///tmp/hive_class_1/AgentPerformance.csv' into table Agent_Performance;
Loading data to table mini_project_1.agent_performance
Table mini_project_1.agent_performance stats: [numFiles=1, totalSize=112661]
OK
Time taken: 0.516 seconds
```

hive> load data local inpath 'file:///tmp/hive_class_1/AgentLogingReport.csv' into table AgentLogingReport;
Loading data to table mini_project_1.agentlogingreport
Table mini_project_1.agentlogingreport stats: [numFiles=1, totalSize=54911]
OK
Time taken: 0.293 seconds

```
hive> load data local inpath 'file:///tmp/hive_class_1/AgentLogingReport.csv' into table AgentLogingReport;
Loading data to table mini_project_1.agentlogingreport
Table mini_project_1.agentlogingreport stats: [numFiles=1, totalSize=54911]
OK
Time taken: 0.293 seconds
```

3.List out all the Agent Names.

hive> select distinct(Agent_name) from Agent_Performance;

```
hive> select distinct(Agent_name) from Agent_Performance;
Query ID = cloudera_20221006190606_c2bf6e85-3d5d-4cdd-b412-e7602ebafc12
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_1662892505239_0133, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1662892505239_0133/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1662892505239_0133
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-06 19:06:47,470 Stage-1 map = 0%,  reduce = 0%
2022-10-06 19:06:55,538 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.95 sec
2022-10-06 19:07:04,347 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 3.51 sec
MapReduce Total cumulative CPU time: 3 seconds 510 msec
Ended Job = job_1662892505239_0133
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 3.51 sec   HDFS Read: 121117 HDFS Write: 867 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 510 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
Jayant Kumar
Jaydeep Dixit
```

hive> select count(distinct(Agent_name)) as agents_count from Agent_Performance;

```
hive> select count(distinct(Agent_name)) as agents_count from Agent_Performance;
Query ID = cloudera_20221006191111_f3e36fb0-3b94-4715-a85e-bb5caf773fe5
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_1662892505239_0135, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1662892505239_0135/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1662892505239_0135
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-06 19:11:27,762 Stage-1 map = 0%,  reduce = 0%
2022-10-06 19:11:33,834 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.46 sec
2022-10-06 19:11:42,551 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 3.35 sec
MapReduce Total cumulative CPU time: 3 seconds 350 msec
Ended Job = job_1662892505239_0135
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 3.35 sec   HDFS Read: 121800 HDFS Write: 3 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 350 msec
OK
agents_count
70
Time taken: 23.705 seconds, Fetched: 1 row(s)
```

4.Find out agent average rating.

hive> select agent_name,avg(average_rating) as average_ratings from Agent_Performance group by agent_name limit 5;

```
hive> select agent_name,avg(average_rating) as average_ratings from Agent_Performance group by agent_name limit 5;
Query ID = cloudera_20221006191818_8f5a622d-f1e6-4942-87f1-11c6f3b764b8
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_1662892505239_0138, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1662892505239_0138/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1662892505239_0138
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-06 19:18:16,711 Stage-1 map = 0%,  reduce = 0%
2022-10-06 19:18:22,295 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.89 sec
2022-10-06 19:18:30,104 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 3.32 sec
MapReduce Total cumulative CPU time: 3 seconds 320 msec
Ended Job = job_1662892505239_0138
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 3.32 sec   HDFS Read: 122538 HDFS Write: 102 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 320 msec
OK
agent_name      average_ratings
Abhishek        0.0
Aditya  0.0
Aditya Shinde   1.8003333409627278
Aditya_iot      2.3453333377838135
Amersh  0.0
Time taken: 23.906 seconds, Fetched: 5 row(s)
```

5.Total Working days for each agent.

hive> select agent_name,count(distinct(date)) as Total_working_days from AgentLogingReport group by agent_name limit 5;

```
hive> select agent_name,count(distinct(date)) as Total_working_days from AgentLogingReport group by agent_name limit 5;
Query ID = cloudera_20221006192424_60357108-c3e7-4608-9a16-0f01c0a8f6fd
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_1662892505239_0140, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1662892505239_0140/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1662892505239_0140
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-06 19:24:43,800 Stage-1 map = 0%,  reduce = 0%
2022-10-06 19:24:50,469 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.59 sec
2022-10-06 19:24:55,858 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 3.1 sec
MapReduce Total cumulative CPU time: 3 seconds 100 msec
Ended Job = job_1662892505239_0140
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 3.1 sec   HDFS Read: 63905 HDFS Write: 62 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 100 msec
OK
agent_name      total_working_days
Aditya Shinde   1
Aditya_iot      8
Amersh  2
Ameya Jain      7
Ankitjha        2
Time taken: 19.322 seconds, Fetched: 5 row(s)
```

6.Total Query that each agent has taken.

hive> select agent_name,sum(total_chats) as total_query from Agent_performance group by agent_name limit 5;

```
hive> select agent_name,sum(total_chats) as total_query from Agent_performance group by agent_name limit 5;
Query ID = cloudera_20221006192828_c3474d7c-e84d-420f-bba3-2574e700f0c5
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_1662892505239_0141, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1662892505239_0141/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1662892505239_0141
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-06 19:28:55,350 Stage-1 map = 0%,  reduce = 0%
2022-10-06 19:29:00,750 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.61 sec
2022-10-06 19:29:06,972 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 2.85 sec
MapReduce Total cumulative CPU time: 2 seconds 850 msec
Ended Job = job_1662892505239_0141
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 2.85 sec   HDFS Read: 122052 HDFS Write: 66 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 850 msec
OK
agent_name      total_query
Abhishek        0
Aditya  0
Aditya Shinde   277
Aditya_iot      231
Amersh  0
Time taken: 19.972 seconds, Fetched: 5 row(s)
```

7.Total feedback that each agent has received.

hive> select agent_name,sum(total_feedback) as total_feedback_received from Agent_Performance group by agent_name limit 5;

```
hive> select agent_name,sum(total_feedback) as total_feedback_received from Agent_Performance group by agent_name limit 5;
Query ID = cloudera_20221006193232_67c8f2c0-0476-4d2d-9115-8756fdbe61ab
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_1662892505239_0142, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1662892505239_0142/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1662892505239_0142
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-06 19:33:02,617 Stage-1 map = 0%,  reduce = 0%
2022-10-06 19:33:07,861 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.46 sec
2022-10-06 19:33:14,079 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 2.73 sec
MapReduce Total cumulative CPU time: 2 seconds 730 msec
Ended Job = job_1662892505239_0142
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 2.73 sec   HDFS Read: 122070 HDFS Write: 66 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 730 msec
OK
agent_name      total_feedback_received
Abhishek        0
Aditya  0
Aditya Shinde   153
Aditya_iot      131
Amersh  0
Time taken: 18.705 seconds, Fetched: 5 row(s)
```

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

hive> select agent_name,average_rating from Agent_performance where average_rating between 3.5 and 4 limit 5;

```
hive> select agent_name,average_rating from Agent_performance where average_rating between 3.5 and 4 limit 5;
OK
agent_name      average_rating
Swati   3.67
Manjunatha A    3.6
Boktiar Ahmed Bappy     4.0
Prateek _iot    3.75
Nandani Gupta   3.79
Time taken: 0.059 seconds, Fetched: 5 row(s)
hive>
```

9.Agent name who have average rating less than 3.5.

hive> select agent_name,average_rating from Agent_performance where average_rating<3.5 limit 5;

```
hive> select agent_name,average_rating from Agent_performance where average_rating<3.5 limit 5;
OK
agent_name      average_rating
Nandani Gupta   3.14
Hitesh Choudhary        0.0
Sanjeevan       0.0
Anirudh         0.0
Shiva Srivastava        0.0
Time taken: 0.075 seconds, Fetched: 5 row(s)
hive>
```

10.Agent name who have average rating more than 4.5

hive> select agent_name,average_rating from Agent_performance where average_rating>4.5 limit 5;

```
hive> select agent_name,average_rating from Agent_performance where average_rating>4.5 limit 5;
OK
agent_name      average_rating
Ameya Jain      4.55
Mahesh Sarade   4.71
Mukesh  4.62
Saikumarreddy N 5.0
Sanjeev Kumar   5.0
Time taken: 0.05 seconds, Fetched: 5 row(s)
hive>
```

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

hive> select agent_name, avg(Total_feedback) from Agent_performance group by agent_name having avg(Total_feedback)>4.5 limit 5;

```
hive> select agent_name, avg(Total_feedback) from Agent_performance group by agent_name having avg(Total_feedback)>4.5 limit 5;
Query ID = cloudera_20221006195252_6f69077f-b77c-4913-b66e-f0d606ba3307
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_1662892505239_0147, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1662892505239_0147/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1662892505239_0147
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-06 19:52:32,323 Stage-1 map = 0%,  reduce = 0%
2022-10-06 19:52:37,747 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.51 sec
2022-10-06 19:52:45,014 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 3.14 sec
MapReduce Total cumulative CPU time: 3 seconds 140 msec
Ended Job = job_1662892505239_0147
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 3.14 sec   HDFS Read: 122940 HDFS Write: 120 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 140 msec
OK
agent_name      _c1
Aditya Shinde   5.1
Ameya Jain      7.6
Aravind         7.766666666666667
Ayushi Mishra   10.966666666666667
Bharath         8.233333333333333
Time taken: 20.119 seconds, Fetched: 5 row(s)
```

12.Average weekly response time for each agent.

hive> select s.agent_name,avg(total_time[0]*3600+total_time[1]*60+total_time[2])/3600 as Average_time from (select agent_name,split(average_response_time,':') as total_time from Agent_performance) s group by s.agent_name limit 5;

```
hive> select s.agent_name,avg(total_time[0]*3600+total_time[1]*60+total_time[2])/3600 as Average_time from (select agent_name,split(average_response_time,':') as total_time from Agent_perfo
rmance) s group by s.agent_name limit 5;
Query ID = cloudera_20221006201414_0a30eab8-99a5-4b06-8680-266a1ca50534
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_1662892505239_0152, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1662892505239_0152/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1662892505239_0152
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-06 20:14:09,346 Stage-1 map = 0%,  reduce = 0%
2022-10-06 20:14:15,640 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 2.51 sec
2022-10-06 20:14:22,909 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 4.31 sec
MapReduce Total cumulative CPU time: 4 seconds 310 msec
Ended Job = job_1662892505239_0152
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 4.31 sec   HDFS Read: 125412 HDFS Write: 105 SUCCESS
Total MapReduce CPU Time Spent: 4 seconds 310 msec
OK
s.agent_name    average_time
Abhishek        0.0
Aditya  0.0
Aditya Shinde   0.00825925925925926
Aditya_iot      0.009435185185185185
Amersh  0.0
Time taken: 20.114 seconds, Fetched: 5 row(s)
hive>
```

13.Average weekly resolution time for each agent.

hive> select s.agent_name,avg(total_time[0]*3600+total_time[1]*60+total_time[2])/3600 as Average_time from (select agent_name,split(average_resolution_time,':') as total_time from Agent_performance) s group by s.agent_name limit 5;

```
hive> select s.agent_name,avg(total_time[0]*3600+total_time[1]*60+total_time[2])/3600 as Average_time from (select agent_name,split(average_resolution_time,':') as total_time from Agent_per
formance) s group by s.agent_name limit 5;
Query ID = cloudera_20221006201717_9efa12ef-c017-42db-8c6a-f795183a5f7f
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_1662892505239_0153, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1662892505239_0153/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1662892505239_0153
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-06 20:17:26,930 Stage-1 map = 0%,  reduce = 0%
2022-10-06 20:17:32,338 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.87 sec
2022-10-06 20:17:38,601 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 3.39 sec
MapReduce Total cumulative CPU time: 3 seconds 390 msec
Ended Job = job_1662892505239_0153
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 3.39 sec   HDFS Read: 125412 HDFS Write: 104 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 390 msec
OK
s.agent_name    average_time
Abhishek        0.0
Aditya  0.0
Aditya Shinde   0.17239814814814813
Aditya_iot      0.16369444444444442
Amersh  0.0
Time taken: 19.073 seconds, Fetched: 5 row(s)
hive>
```

14. Find the number of chats on which they have received feedback.
hive> select agent_name,sum(total_chats) as sum_of_chats,total_feedback from Agent_Performance group by agent_name,total_feedback having total_feedback>0 limit 5;

```
hive> select agent_name,sum(total_chats) as sum_of_chats,total_feedback from Agent_Performance group by agent_name,total_feedback having total_feedback>0 limit 5;
Query ID = cloudera_20221006202626_cf4c3e15-f84a-4ee4-a5ed-9a3243339132
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_1662892505239_0156, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1662892505239_0156/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1662892505239_0156
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-06 20:26:48,247 Stage-1 map = 0%,  reduce = 0%
2022-10-06 20:26:53,483 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.57 sec
2022-10-06 20:26:59,742 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 3.25 sec
MapReduce Total cumulative CPU time: 3 seconds 250 msec
Ended Job = job_1662892505239_0156
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 3.25 sec   HDFS Read: 122681 HDFS Write: 96 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 250 msec
OK
agent_name      sum_of_chats    total_feedback
Aditya Shinde   8       7
Aditya Shinde   17      8
Aditya Shinde   67      9
Aditya Shinde   18      11
Aditya Shinde   27      14
Time taken: 19.517 seconds, Fetched: 5 row(s)
```

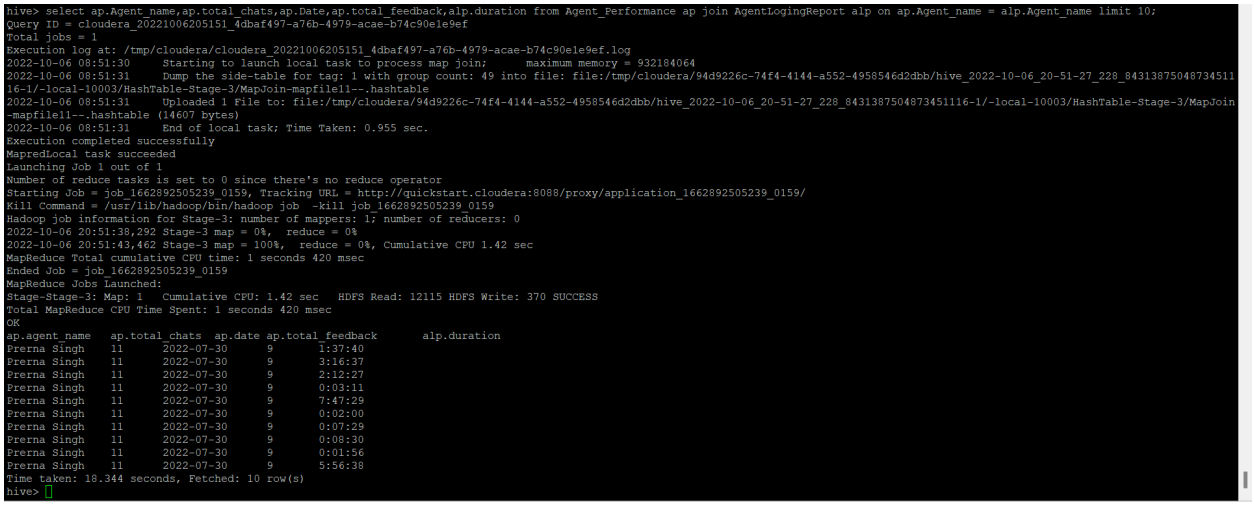
15.Total Contribution hour for each and every agent's weekly basis.

hive> select agent_name,sum(total_time[0]*3600+total_time[1]*60+total_time[2])/3600 Hours_time,week from (select agent_name,split(duration,':') as total_time, weekofyear(Date) as week from AgentLogingReport) s group by agent_name,week limit 5;

```
hive> select agent_name,sum(total_time[0]*3600+total_time[1]*60+total_time[2])/3600 Hours_time,week from (select agent_name,split(duration,':') as total_time, weekofyear(Date) as week from
AgentLogingReport) s group by agent_name,week limit 5;
Query ID = cloudera_20221006203636_8e1bc346-dbdb-431c-be20-4aa489e7ea41
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_1662892505239_0157, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1662892505239_0157/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1662892505239_0157
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-06 20:36:17,651 Stage-1 map = 0%,  reduce = 0%
2022-10-06 20:36:23,860 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.7 sec
2022-10-06 20:36:29,082 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 3.53 sec
MapReduce Total cumulative CPU time: 3 seconds 530 msec
Ended Job = job_1662892505239_0157
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 3.53 sec   HDFS Read: 67293 HDFS Write: 162 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 530 msec
OK
agent_name      hours_time      week
Aditya Shinde   0.03611111111111111     30
Aditya_iot      6.095277777777778       29
Aditya_iot      9.63583333333334        30
Amersh  3.063888888888889       30
Ameya Jain      24.083055555555557      29
Time taken: 19.204 seconds, Fetched: 5 row(s)
hive>
```

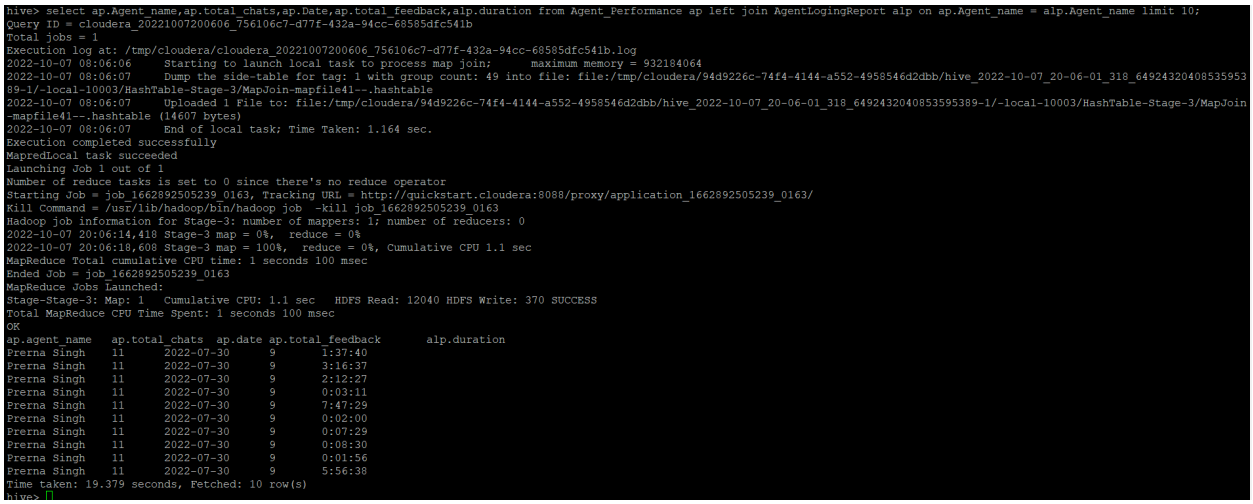
16.Perform inner join, left join and right join based on agent_name column and after joining the tables export the data into your local system.

Inner Join:-
hive> select ap.Agent_name,ap.total_chats,ap.Date,ap.total_feedback,alp.duration from Agent_Performance ap join AgentLogingReport alp on ap.Agent_name = alp.Agent_name limit 10;



Left Join:-

hive> select ap.Agent_name,ap.total_chats,ap.Date,ap.total_feedback,alp.duration from Agent_Performance ap left join AgentLogingReport alp on ap.Agent_name = alp.Agent_name limit 10;



Right Join:-

hive> select ap.Agent_name,ap.total_chats,ap.Date,ap.total_feedback,alp.duration from Agent_Performance ap right join AgentLogingReport alp on ap.Agent_name = alp.Agent_name limit 10;

```
hive> select ap.Agent_name,ap.total_chats,ap.Date,ap.total_feedback,alp.duration from Agent_Performance ap right join AgentLogingReport alp on ap.Agent_name = alp.Agent_name limit 10;
Query ID = cloudera_20221007200707_a1152b82-ba0a-40b5-8aa4-5f1526c94a5b
Total jobs = 1
Execution log at: /tmp/cloudera/cloudera_20221007200707_a1152b82-ba0a-40b5-8aa4-5f1526c94a5b.log
2022-10-07 08:07:48     Starting to launch local task to process map join;      maximum memory = 932184064
2022-10-07 08:07:49     Dump the side-table for tag: 0 with group count: 70 into file: file:/tmp/cloudera/94d9226c-74f4-4144-a552-4958546d2dbb/hive_2022-10-07_20-07-45_999_36503607748180929
32-1/-local-10003/HashTable-Stage-3/MapJoin-mapfile50--.hashtable
2022-10-07 08:07:49     Uploaded 1 File to: file:/tmp/cloudera/94d9226c-74f4-4144-a552-4958546d2dbb/hive_2022-10-07_20-07-45_999_3650360774818092932-1/-local-10003/HashTable-Stage-3/MapJoin
-mapfile50--.hashtable (23750 bytes)
2022-10-07 08:07:49     End of local task; Time Taken: 0.842 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_1662892505239_0164, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1662892505239_0164/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1662892505239_0164
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0
2022-10-07 20:07:54,725 Stage-3 map = 0%,  reduce = 0%
2022-10-07 20:07:59,868 Stage-3 map = 100%,  reduce = 0%, Cumulative CPU 1.14 sec
MapReduce Total cumulative CPU time: 1 seconds 140 msec
Ended Job = job_1662892505239_0164
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1   Cumulative CPU: 1.14 sec   HDFS Read: 11892 HDFS Write: 429 SUCCESS
Total MapReduce CPU Time Spent: 1 seconds 140 msec
OK
ap.agent_name   ap.total_chats  ap.date ap.total_feedback       alp.duration
Shivananda Sonwane      4       2022-07-30      1       2:04:10
Shivananda Sonwane      14      2022-07-29      9       2:04:10
Shivananda Sonwane      5       2022-07-28      4       2:04:10
Shivananda Sonwane      26      2022-07-27      18      2:04:10
Shivananda Sonwane      24      2022-07-26      14      2:04:10
Shivananda Sonwane      13      2022-07-25      8       2:04:10
Shivananda Sonwane      8       2022-07-24      5       2:04:10
Shivananda Sonwane      20      2022-07-23      13      2:04:10
Shivananda Sonwane      16      2022-07-22      9       2:04:10
Shivananda Sonwane      7       2022-07-21      6       2:04:10
Time taken: 14.923 seconds, Fetched: 10 row(s)
hive>
```

Exporting Data to local:-

Inner-Join:-

Query in inner_join.hql

```
cloudera@quickstart:~
select ap.Agent_name,ap.total_chats,ap.Date,ap.total_feedback,alp.duration from mini_project_1.Agent_Performance ap join mini_project_1.AgentLogingReport alp on ap.Agent_name = alp.Agent_na
me;
~
~
~
```

→hive -f inner_join.hql >> inner_join.csv

```
[cloudera@quickstart ~]$ vi inner_join.hql
[cloudera@quickstart ~]$ hive -f inner_join.hql >> inner_join.csv

Logging initialized using configuration in file:/etc/hive/conf.dist/hive-log4j.properties
Query ID = cloudera_20221007200303_003474cb-9443-4e56-a254-11777961b895
Total jobs = 1
Execution log at: /tmp/cloudera/cloudera_20221007200303_003474cb-9443-4e56-a254-11777961b895.log
2022-10-07 08:03:35     Starting to launch local task to process map join;      maximum memory = 932184064
2022-10-07 08:03:36     Dump the side-table for tag: 1 with group count: 49 into file: file:/tmp/cloudera/72abecf4-55e1-4f01-89cd-521efb5a1873/hive_2022-10-07_20-03-31_197_12756355911059824
23-1/-local-10003/HashTable-Stage-3/MapJoin-mapfile01--.hashtable
2022-10-07 08:03:36     Uploaded 1 File to: file:/tmp/cloudera/72abecf4-55e1-4f01-89cd-521efb5a1873/hive_2022-10-07_20-03-31_197_1275635591105982423-1/-local-10003/HashTable-Stage-3/MapJoin
-mapfile01--.hashtable (14607 bytes)
2022-10-07 08:03:36     End of local task; Time Taken: 1.359 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_1662892505239_0162, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1662892505239_0162/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1662892505239_0162
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0
2022-10-07 20:03:44,048 Stage-3 map = 0%,  reduce = 0%
2022-10-07 20:03:53,733 Stage-3 map = 100%,  reduce = 0%, Cumulative CPU 3.65 sec
MapReduce Total cumulative CPU time: 3 seconds 650 msec
Ended Job = job_1662892505239_0162
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1   Cumulative CPU: 3.65 sec   HDFS Read: 120418 HDFS Write: 823577 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 650 msec
OK
Time taken: 24.769 seconds, Fetched: 21180 row(s)
[cloudera@quickstart ~]$ ls
bucket_data.txt     cloudera-manager  Documents  employee.csv              inner_join.csv  kerberos       many_udf.py      Music     Public          two.txt
bucket_location.txt  cm_api.py        Downloads  enterprise-deployment.json  inner_join.hql  lib            many_udfs.py     parcels   sales_data.csv  Videos
chinna              Desktop           eclipse    express-deployment.json     json_file.json  locations.csv  multiply_udf.py  Pictures  Templates       workspace
[cloudera@quickstart ~]$ wc -l inner_join.csv
21184 inner_join.csv
```

Left Join:-

Query in left_join.hql

```
cloudera@quickstart:~
select ap.Agent_name,ap.total_chats,ap.Date,ap.total_feedback,alp.duration from mini_project_1.Agent_Performance ap left join mini_project_1.AgentLogingReport alp on ap.Agent_name = alp.Age
nt_name;
~
~
```

→ hive -f left_join.hql >> left_join.csv

```
[cloudera@quickstart ~]$ vi left_join.hql
[cloudera@quickstart ~]$ hive -f left_join.hql >> left_join.csv

Logging initialized using configuration in file:/etc/hive/conf.dist/hive-log4j.properties
Query ID = cloudera_20221007201717_f01df82c-2756-4818-81fe-864da022334f
Total jobs = 1
Execution log at: /tmp/cloudera/cloudera_20221007201717_f01df82c-2756-4818-81fe-864da022334f.log
2022-10-07 08:18:03     Starting to launch local task to process map join;      maximum memory = 932184064
2022-10-07 08:18:03     Dump the side-table for tag: 1 with group count: 49 into file: file:/tmp/cloudera/29b4fb38-6dd8-4d94-b134-2edab7328464/hive_2022-10-07_20-17-59_410_8198093207034877018-1/-local-10003/HashTable-Stage-3/MapJoin-mapfile01--.hashtable
2022-10-07 08:18:03     Uploaded 1 File to: file:/tmp/cloudera/29b4fb38-6dd8-4d94-b134-2edab7328464/hive_2022-10-07_20-17-59_410_8198093207034877018-1/-local-10003/HashTable-Stage-3/MapJoin-mapfile01--.hashtable (14607 bytes)
2022-10-07 08:18:03     End of local task; Time Taken: 0.739 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_1662892505239_0165, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1662892505239_0165/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1662892505239_0165
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0
2022-10-07 20:18:09,975 Stage-3 map = 0%,  reduce = 0%
2022-10-07 20:18:18,357 Stage-3 map = 100%,  reduce = 0%, Cumulative CPU 2.45 sec
MapReduce Total cumulative CPU time: 2 seconds 450 msec
Ended Job = job_1662892505239_0165
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1   Cumulative CPU: 2.45 sec   HDFS Read: 120275 HDFS Write: 854177 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 450 msec
OK
Time taken: 20.087 seconds, Fetched: 22260 row(s)
[cloudera@quickstart ~]$ ls
bucket_data.txt      cloudera-manager  Documents  employee.csv              inner_join.csv  kerberos       lib            many_udfs.py     parcels   sales_data.csv  Videos
bucket_location.txt  cm_api.py         Downloads  enterprise-deployment.json  inner_join.hql  left_join.csv  locations.csv  multiply_udf.py  Pictures  Templates       workspace
chinna               Desktop           eclipse    express-deployment.json     json_file.json  left_join.hql  many_udf.py    Music            Public    two.txt
[cloudera@quickstart ~]$ wc -l left_join.csv
22264 left_join.csv
[cloudera@quickstart ~]$
```

Right Join:-

Query in Right_join

```
cloudera@quickstart:~
select ap.Agent_name,ap.total_chats,ap.Date,ap.total_feedback,alp.duration from mini_project_1.Agent_Performance ap right join mini_project_1.AgentLogingReport alp on ap.Agent_name = alp.Agent_name;
~
~
```

→ hive -f right_join.hql >> right_join.csv

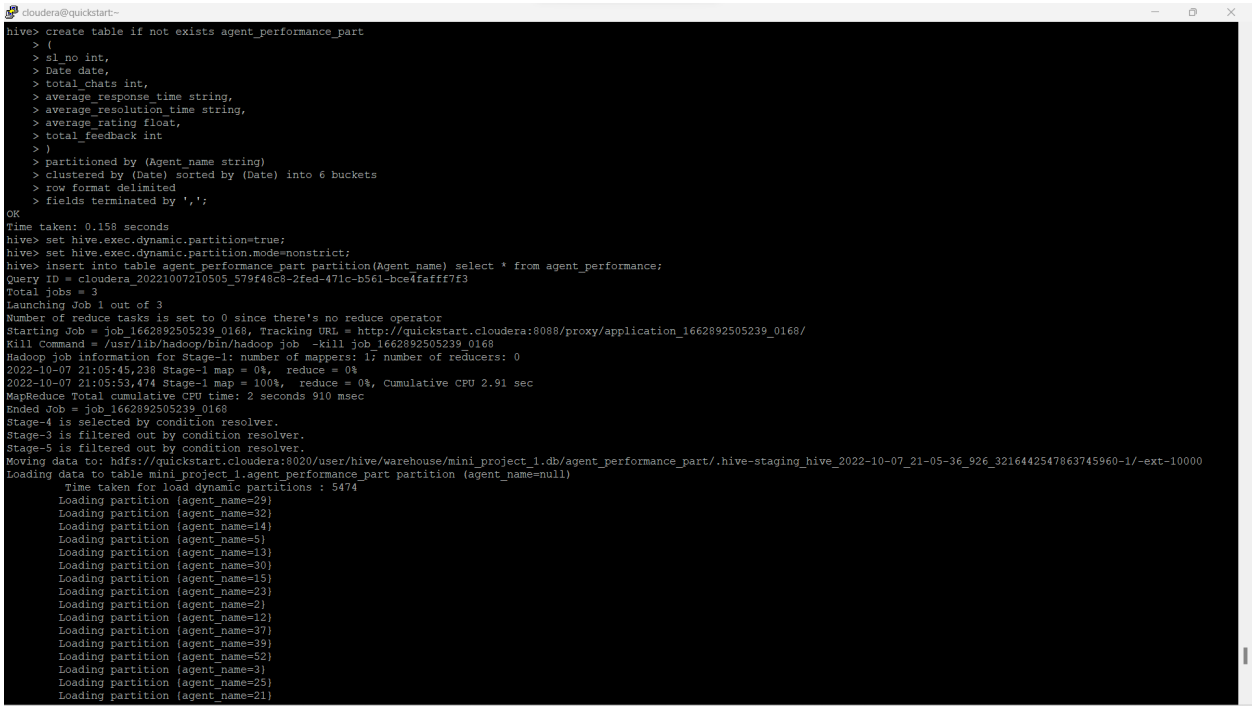
```
[cloudera@quickstart ~]$ vi right_join.hql
[cloudera@quickstart ~]$ hive -f right_join.hql >> right_join.csv

Logging initialized using configuration in file:/etc/hive/conf.dist/hive-log4j.properties
Query ID = cloudera_20221007202424_04d3a3af-8b54-4f9a-a3ed-10da81165e8b
Total jobs = 1
Execution log at: /tmp/cloudera/cloudera_20221007202424_04d3a3af-8b54-4f9a-a3ed-10da81165e8b.log
2022-10-07 08:24:29     Starting to launch local task to process map join;      maximum memory = 932184064
2022-10-07 08:24:30     Dump the side-table for tag: 0 with group count: 70 into file: file:/tmp/cloudera/fb4f3de4-40be-4a28-948c-9dff6a24fdaf/hive_2022-10-07_20-24-25_253_1082539476716018447-1/-local-10003/HashTable-Stage-3/MapJoin-mapfile00--.hashtable
2022-10-07 08:24:30     Uploaded 1 File to: file:/tmp/cloudera/fb4f3de4-40be-4a28-948c-9dff6a24fdaf/hive_2022-10-07_20-24-25_253_1082539476716018447-1/-local-10003/HashTable-Stage-3/MapJoin-mapfile00--.hashtable (23750 bytes)
2022-10-07 08:24:30     End of local task; Time Taken: 0.844 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_1662892505239_0166, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1662892505239_0166/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1662892505239_0166
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0
2022-10-07 20:24:37,645 Stage-3 map = 0%,  reduce = 0%
2022-10-07 20:24:44,109 Stage-3 map = 100%,  reduce = 0%, Cumulative CPU 2.66 sec
MapReduce Total cumulative CPU time: 2 seconds 660 msec
Ended Job = job_1662892505239_0166
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1   Cumulative CPU: 2.66 sec   HDFS Read: 62376 HDFS Write: 829699 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 660 msec
OK
Time taken: 21.056 seconds, Fetched: 21486 row(s)
[cloudera@quickstart ~]$ ls
bucket_data.txt      cloudera-manager  Documents  employee.csv              inner_join.csv  kerberos       lib            many_udfs.py     parcels   right_join.csv  Templates  workspace
bucket_location.txt  cm_api.py         Downloads  enterprise-deployment.json  inner_join.hql  left_join.csv  locations.csv  multiply_udf.py  Pictures  right_join.hql  two.txt
chinna               Desktop           eclipse    express-deployment.json     json_file.json  left_join.hql  many_udf.py    Music            Public    sales_data.csv  Videos
[cloudera@quickstart ~]$ wc -l right_join.csv
21488 right_join.csv
[cloudera@quickstart ~]$
```

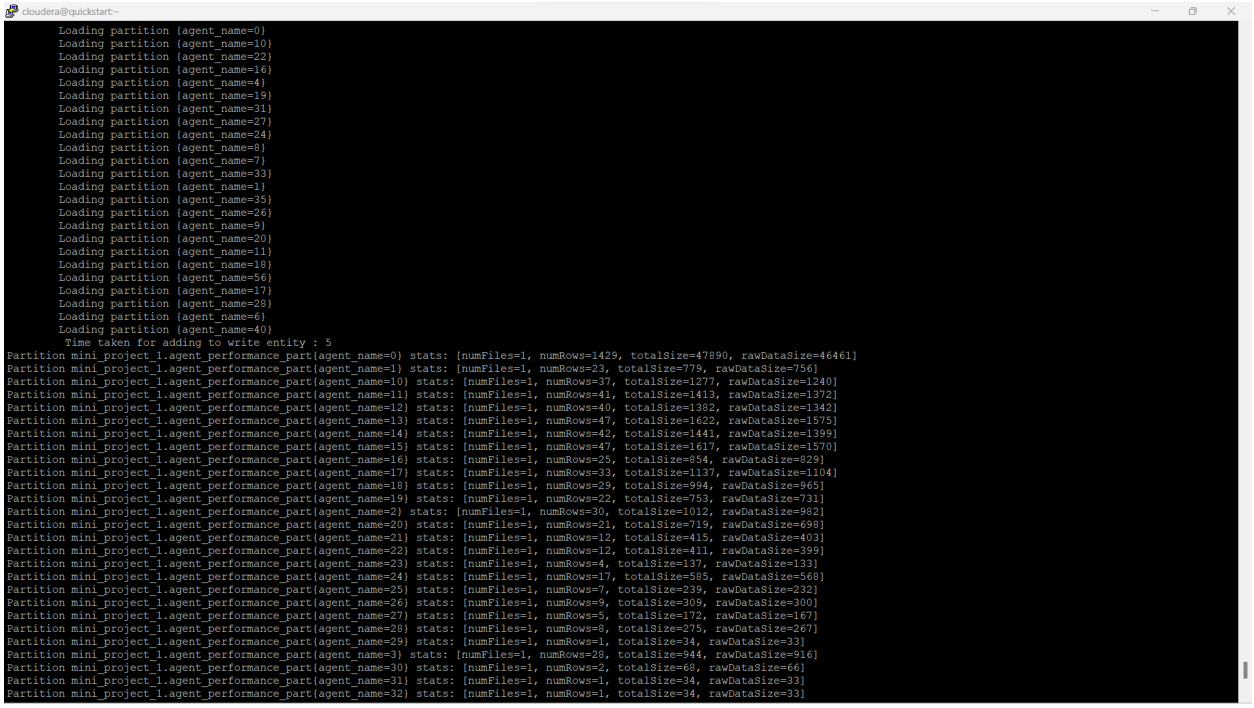
17. Perform partitioning on top of the agent column and then on top of that perform bucketing for each partitioning.

For agent_performance table:-

```
hive> create table if not exists agent_performance_part
    > (
    > sl_no int,
    > Date date,
    > total_chats int,
    > average_response_time string,
```

```
    > average_resolution_time string,
    > average_rating float,
    > total_feedback int
    > )
    > partitioned by (Agent_name string)
    > clustered by (Date) sorted by (Date) into 6 buckets
    > row format delimited
    > fields terminated by ',';
OK
Time taken: 0.158 seconds
hive> set hive.exec.dynamic.partition=true;
hive> set hive.exec.dynamic.partition.mode=nonstrict;
hive> insert into table agent_performance_part partition(Agent_name) select * from
agent_performance;
```



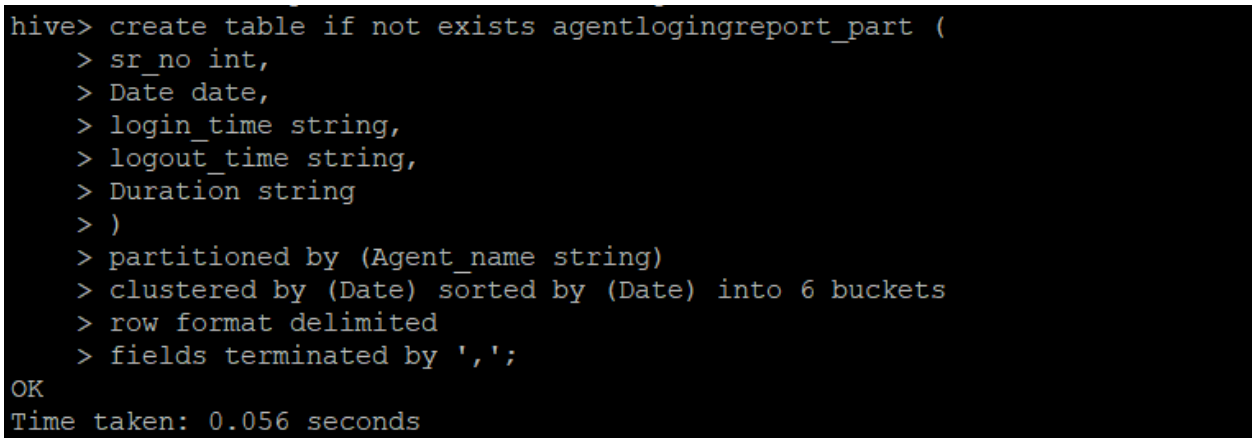


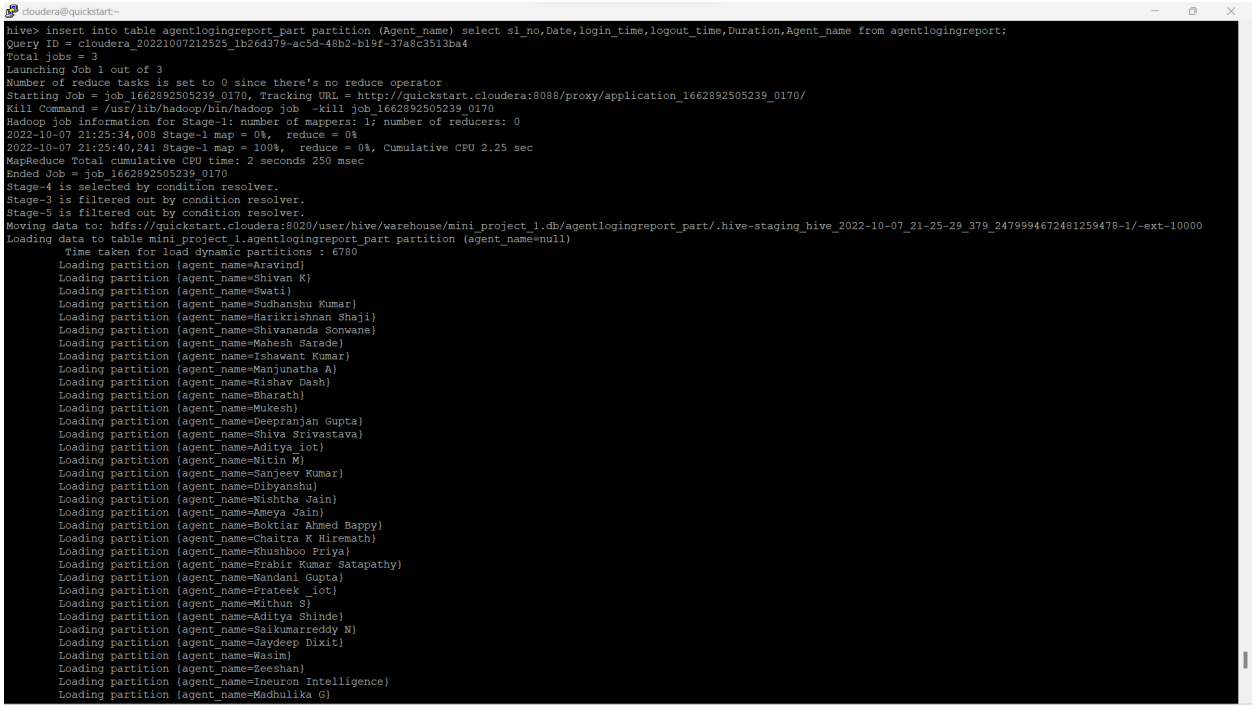

```
Partition mini_project_1.agent_performance_part{agent_name=31} stats: [numFiles=1, numRows=1, totalSize=34, rawDataSize=33]
Partition mini_project_1.agent_performance_part{agent_name=32} stats: [numFiles=1, numRows=1, totalSize=34, rawDataSize=33]
Partition mini_project_1.agent_performance_part{agent_name=33} stats: [numFiles=1, numRows=3, totalSize=103, rawDataSize=100]
Partition mini_project_1.agent_performance_part{agent_name=35} stats: [numFiles=1, numRows=2, totalSize=69, rawDataSize=67]
Partition mini_project_1.agent_performance_part{agent_name=37} stats: [numFiles=1, numRows=3, totalSize=104, rawDataSize=101]
Partition mini_project_1.agent_performance_part{agent_name=39} stats: [numFiles=1, numRows=1, totalSize=35, rawDataSize=34]
Partition mini_project_1.agent_performance_part{agent_name=4} stats: [numFiles=1, numRows=29, totalSize=975, rawDataSize=946]
Partition mini_project_1.agent_performance_part{agent_name=40} stats: [numFiles=1, numRows=1, totalSize=34, rawDataSize=33]
Partition mini_project_1.agent_performance_part{agent_name=5} stats: [numFiles=1, numRows=24, totalSize=816, rawDataSize=792]
Partition mini_project_1.agent_performance_part{agent_name=52} stats: [numFiles=1, numRows=1, totalSize=35, rawDataSize=34]
Partition mini_project_1.agent_performance_part{agent_name=56} stats: [numFiles=1, numRows=1, totalSize=34, rawDataSize=33]
Partition mini_project_1.agent_performance_part{agent_name=6} stats: [numFiles=1, numRows=32, totalSize=1090, rawDataSize=1058]
Partition mini_project_1.agent_performance_part{agent_name=7} stats: [numFiles=1, numRows=26, totalSize=888, rawDataSize=862]
Partition mini_project_1.agent_performance_part{agent_name=8} stats: [numFiles=1, numRows=23, totalSize=787, rawDataSize=764]
Partition mini_project_1.agent_performance_part{agent_name=9} stats: [numFiles=1, numRows=41, totalSize=1411, rawDataSize=1370]
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1   Cumulative CPU: 2.91 sec   HDFS Read: 118830 HDFS Write: 75799 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 910 msec
OK
_col0   _col1   _col2   _col3   _col4   _col5   _col6   _col7
Time taken: 24.795 seconds
```

For AgentLogingReport:-

```
hive> create table if not exists agentlogingreport_part (
    > sr_no int,
    > Date date,
    > login_time string,
    > logout_time string,
    > Duration string
    > )
    > partitioned by (Agent_name string)
    > clustered by (Date) sorted by (Date) into 6 buckets
```

```
    > row format delimited
    > fields terminated by ',';
OK
Time taken: 0.056 seconds
hive> set hive.exec.dynamic.partition=true;
hive> set hive.exec.dynamic.partition.mode=nonstrict;
hive> insert into table agentlogingreport_part partition (Agent_name) select
sl_no,Date,login_time,logout_time,Duration,Agent_name from agentlogingreport;
```





cloudera@quickstart:~

```
        Loading partition {agent_name=Muskan Garg}
        Loading partition {agent_name=Saurabh Shukla}
        Loading partition {agent_name=Shubham Sharma}
        Loading partition {agent_name=Prerna Singh}
        Loading partition {agent_name=Suraj S Bilgi}
        Loading partition {agent_name=Maitry}
        Loading partition {agent_name=Hrisikesh Neogi}
        Loading partition {agent_name=Anurag Tiwari}
        Loading partition {agent_name=Tarun}
        Loading partition {agent_name=Sowmiya Sivakumar}
        Loading partition {agent_name=Ayushi Mishra}
        Loading partition {agent_name=Hyder Abbas}
        Loading partition {agent_name=Jawala Prakash}
        Loading partition {agent_name=Amersh}
        Loading partition {agent_name=Ankitjha}
          Time taken for adding to write entity : 5
Partition mini_project_1.agentlogingreport_part{agent_name=Aditya Shinde} stats: [numFiles=1, numRows=1, totalSize=41, rawDataSize=40]
Partition mini_project_1.agentlogingreport_part{agent_name=Aditya_iot} stats: [numFiles=1, numRows=9, totalSize=368, rawDataSize=359]
Partition mini_project_1.agentlogingreport_part{agent_name=Amersh} stats: [numFiles=1, numRows=4, totalSize=164, rawDataSize=160]
Partition mini_project_1.agentlogingreport_part{agent_name=Ameya Jain} stats: [numFiles=1, numRows=10, totalSize=409, rawDataSize=399]
Partition mini_project_1.agentlogingreport_part{agent_name=Ankitjha} stats: [numFiles=1, numRows=4, totalSize=164, rawDataSize=160]
Partition mini_project_1.agentlogingreport_part{agent_name=Anurag Tiwari} stats: [numFiles=1, numRows=37, totalSize=1500, rawDataSize=1463]
Partition mini_project_1.agentlogingreport_part{agent_name=Aravind} stats: [numFiles=1, numRows=10, totalSize=409, rawDataSize=399]
Partition mini_project_1.agentlogingreport_part{agent_name=Ayushi Mishra} stats: [numFiles=1, numRows=18, totalSize=729, rawDataSize=711]
Partition mini_project_1.agentlogingreport_part{agent_name=Bharath} stats: [numFiles=1, numRows=9, totalSize=359, rawDataSize=350]
Partition mini_project_1.agentlogingreport_part{agent_name=Boktiar Ahmed Bappy} stats: [numFiles=1, numRows=17, totalSize=680, rawDataSize=663]
Partition mini_project_1.agentlogingreport_part{agent_name=Chaitra K Hiremath} stats: [numFiles=1, numRows=13, totalSize=530, rawDataSize=517]
Partition mini_project_1.agentlogingreport_part{agent_name=Deepranjan Gupta} stats: [numFiles=1, numRows=58, totalSize=2359, rawDataSize=2301]
Partition mini_project_1.agentlogingreport_part{agent_name=Dibyanshu} stats: [numFiles=1, numRows=208, totalSize=8361, rawDataSize=8153]
Partition mini_project_1.agentlogingreport_part{agent_name=Harikrishnan Shaji} stats: [numFiles=1, numRows=23, totalSize=931, rawDataSize=908]
Partition mini_project_1.agentlogingreport_part{agent_name=Hrisikesh Neogi} stats: [numFiles=1, numRows=37, totalSize=1493, rawDataSize=1456]
Partition mini_project_1.agentlogingreport_part{agent_name=Hyder Abbas} stats: [numFiles=1, numRows=2, totalSize=82, rawDataSize=80]
Partition mini_project_1.agentlogingreport_part{agent_name=Ineuron Intelligence} stats: [numFiles=1, numRows=1, totalSize=41, rawDataSize=40]
Partition mini_project_1.agentlogingreport_part{agent_name=Ishawant Kumar} stats: [numFiles=1, numRows=49, totalSize=1980, rawDataSize=1931]
Partition mini_project_1.agentlogingreport_part{agent_name=Jawala Prakash} stats: [numFiles=1, numRows=16, totalSize=636, rawDataSize=620]
Partition mini_project_1.agentlogingreport_part{agent_name=Jaydeep Dixit} stats: [numFiles=1, numRows=11, totalSize=449, rawDataSize=438]
Partition mini_project_1.agentlogingreport_part{agent_name=Khushboo Priya} stats: [numFiles=1, numRows=18, totalSize=726, rawDataSize=708]
Partition mini_project_1.agentlogingreport_part{agent_name=Madhulika G} stats: [numFiles=1, numRows=17, totalSize=687, rawDataSize=670]
Partition mini_project_1.agentlogingreport_part{agent_name=Mahesh Sarade} stats: [numFiles=1, numRows=36, totalSize=1468, rawDataSize=1432]
Partition mini_project_1.agentlogingreport_part{agent_name=Maitry} stats: [numFiles=1, numRows=5, totalSize=200, rawDataSize=195]
Partition mini_project_1.agentlogingreport_part{agent_name=Manjunatha A} stats: [numFiles=1, numRows=8, totalSize=318, rawDataSize=310]
Partition mini_project_1.agentlogingreport_part{agent_name=Mithun S} stats: [numFiles=1, numRows=14, totalSize=564, rawDataSize=550]
Partition mini_project_1.agentlogingreport_part{agent_name=Mukesh} stats: [numFiles=1, numRows=3, totalSize=121, rawDataSize=118]
Partition mini_project_1.agentlogingreport_part{agent_name=Muskan Garg} stats: [numFiles=1, numRows=12, totalSize=491, rawDataSize=479]
Partition mini_project_1.agentlogingreport_part{agent_name=Nandani Gupta} stats: [numFiles=1, numRows=11, totalSize=447, rawDataSize=436]
Partition mini_project_1.agentlogingreport_part{agent_name=Nishtha Jain} stats: [numFiles=1, numRows=18, totalSize=728, rawDataSize=710]
Partition mini_project_1.agentlogingreport_part{agent_name=Nitin M} stats: [numFiles=1, numRows=1, totalSize=41, rawDataSize=40]
Partition mini_project_1.agentlogingreport_part{agent_name=Prabir Kumar Satapathy} stats: [numFiles=1, numRows=26, totalSize=1054, rawDataSize=1028]
Partition mini_project_1.agentlogingreport_part{agent_name=Prateek _iot} stats: [numFiles=1, numRows=17, totalSize=690, rawDataSize=673]
Partition mini_project_1.agentlogingreport_part{agent_name=Prerna Singh} stats: [numFiles=1, numRows=18, totalSize=735, rawDataSize=717]
Partition mini_project_1.agentlogingreport_part{agent_name=Rishav Dash} stats: [numFiles=1, numRows=12, totalSize=481, rawDataSize=469]
Partition mini_project_1.agentlogingreport_part{agent_name=Saikumarreddy N} stats: [numFiles=1, numRows=10, totalSize=410, rawDataSize=400]
```

```
Partition mini_project_1.agentlogingreport_part{agent_name=Sanjeev Kumar} stats: [numFiles=1, numRows=20, totalSize=815, rawDataSize=795]
Partition mini_project_1.agentlogingreport_part{agent_name=Saurabh Shukla} stats: [numFiles=1, numRows=40, totalSize=1608, rawDataSize=1568]
Partition mini_project_1.agentlogingreport_part{agent_name=Shiva Srivastava} stats: [numFiles=1, numRows=15, totalSize=612, rawDataSize=597]
Partition mini_project_1.agentlogingreport_part{agent_name=Shivan K} stats: [numFiles=1, numRows=36, totalSize=1458, rawDataSize=1422]
Partition mini_project_1.agentlogingreport_part{agent_name=Shivananda Sonwane} stats: [numFiles=1, numRows=15, totalSize=600, rawDataSize=585]
Partition mini_project_1.agentlogingreport_part{agent_name=Shubham Sharma} stats: [numFiles=1, numRows=35, totalSize=1425, rawDataSize=1390]
Partition mini_project_1.agentlogingreport_part{agent_name=Sowmiya Sivakumar} stats: [numFiles=1, numRows=24, totalSize=972, rawDataSize=948]
Partition mini_project_1.agentlogingreport_part{agent_name=Sudhanshu Kumar} stats: [numFiles=1, numRows=11, totalSize=446, rawDataSize=435]
Partition mini_project_1.agentlogingreport_part{agent_name=Suraj S Bilgi} stats: [numFiles=1, numRows=5, totalSize=200, rawDataSize=195]
Partition mini_project_1.agentlogingreport_part{agent_name=Swati} stats: [numFiles=1, numRows=5, totalSize=205, rawDataSize=200]
Partition mini_project_1.agentlogingreport_part{agent_name=Tarun} stats: [numFiles=1, numRows=1, totalSize=43, rawDataSize=42]
Partition mini_project_1.agentlogingreport_part{agent_name=Wasim} stats: [numFiles=1, numRows=20, totalSize=813, rawDataSize=793]
Partition mini_project_1.agentlogingreport_part{agent_name=Zeeshan} stats: [numFiles=1, numRows=10, totalSize=409, rawDataSize=399]
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1   Cumulative CPU: 2.25 sec   HDFS Read: 60069 HDFS Write: 44499 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 250 msec
OK
sl_no   date    login_time      logout_time     duration        agent_name
Time taken: 21.402 seconds
hive>
```