

PROJECT DESCRIPTION:- The Project aims to display the Agent logging report and Agent performance data of “I-Neuron technical consultant team”. Both datasets are real-time generated that further contains the details of agents working for the team. The two datasets are stored in the form of CSV files and the project aims to load it to HDFS for further detailed analysis using HIVE.

DESCRIPTION OF DATASET:- The first dataset is Agent_Loading_report dataset which contains almost 1000 records and 7 attributes of the Consultant team agents. The attributes are :- Sln0, Agent,Date,Login_time,Logout_time and duration of activity.

Link- “https://drive.google.com/file/d/1WrG-9qv6atP-W3P_gYln1hHyFKRKMHP/view”

The second dataset is Agent_performance dataset which contains approximately 2200 records and 8 attributes. This dataset mostly depicts the performance analysis of agents working in the technical consultant team. The attributes are:- Sln0,Date,Agent_name,Total_chats,Average_response_time,Average_resolution_time,Average_rating and Total Feedback.

Link - “<https://drive.google.com/file/d/1-JPCZ34dyN6k9CqJa-Y8yxIGq6vTVXU/view>”

PROBLEM STATEMENT:-

There is a list of problem statements which we need to solve for further analysis on the dataset.

- 1.Create a schema based on the given dataset**
- 2. Dump the data inside the hdfs in the given schema location.**

- 3. List of all agents' names.**
- 4. Find out agent average rating.**
- 5. Total working days for each agents**
- 6. Total query that each agent have taken**
- 7. Total Feedback that each agent have received**
- 8. Agent name who have average rating between 3.5 to 4**
- 9. Agent name who have rating less than 3.5**
- 10. Agent name who have rating more than 4.5**
- 11. How many feedback agents have received more than 4.5 average**
- 12. average weekly response time for each agent**
- 13. average weekly resolution time for each agents**
- 14. Find the number of chat on which they have received a feedback**
- 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.**
- 17. Perform partitioning on top of the agent column and then on top of that perform bucketing for each partitioning.**

Hence , the solutions for the above problem statements are shown on the next page. I have performed the analysis on HIVE Engine on Cloudera Platform . The screenshots of the HIVE-Query-language is also pasted below for better understanding.

- 1) **Create a schema based on the given dataset.**
- 2) **Dump the data inside the hdfs in the given schema location.**

First of All, open the Hive Shell in Cloudera terminal and either use the existing database or create a new database. Here, I have used the existing database “hive_assignment” and created an empty table based on the schema of the two datasets.

```
Create table agent_logging_report (  
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");
```

```
Create table agent_performance(  
sl_no int,Date string,Agent_name string,  
total_chats int,average_response_time string,  
average_resolution string,average_rating float,  
total_feedback int)  
row format delimited  
fields terminated by ','  
tblproperties ("skip.header.line.count" = "1");
```

After creating the table, its time to load the csv files in the empty table, At first I have downloaded both the CSV files in my local cloudera manager.

```
Load data local inpath '/home/cloudera/Downloads/  
AgentLoggingReport.csv' into table agent_logging_report;
```

```
Load data local inpath '/home/cloudera/Downloads/  
AgentPerformance.csv' into table agent_performance;
```

The screenshot of the above code part is pasted below.

```

hive> create table agent_logging_report(
  > sno 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: 0.48 seconds
hive> load data local inpath '/home/cloudera/Downloads/AgentLoggingReport.csv' into table agent_logging_report;
Loading data to table hive assignment-agent_logging_report
Table hive assignment-agent_logging_report stats: [numFiles=1, totalSize=53351]
OK
Time taken: 1.077 seconds
hive> select * from agent_logging_report limit 5;
OK
agent_logging_report.sno      agent_logging_report.agent      agent_logging_report.date      agent_logging_report.login_time  agent_logging_report.logout_time agent_logging_report.duration
1      Shivnanda Somane      30-Jul-22      15:35:29      17:39:39      02:04:10
2      Khushboo Priya      30-Jul-22      15:06:59      15:07:16      00:00:17
3      Nandani Gupta      30-Jul-22      15:04:24      17:31:07      02:26:42
4      Harishkesh Neogi      30-Jul-22      14:34:29      15:19:33      00:45:06
5      Mukesh      30-Jul-22      14:03:15      15:11:52      01:08:36
Time taken: 0.183 seconds, Fetched: 5 row(s)
hive> create table agent_performance(
  > sno int,Date string,Agent name string,
  > total_chats int,avg_response_time string,avg_resolution_time string,avg_rating float,
  > total_feedback int)
  > row format delimited
  > fields terminated by ','
  > tblproperties('skip.header.line.count'='1');
OK
Time taken: 0.222 seconds
hive> load data local inpath '/home/cloudera/Downloads/AgentPerformance.csv' into table agent_performance;
Loading data to table hive assignment-agent_performance
Table hive assignment-agent_performance stats: [numFiles=1, totalSize=199853]
OK
Time taken: 0.626 seconds
hive> select * from agent_performance limit 5;
OK
agent_performance.sno      agent_performance.date      agent_performance.agent_name      agent_performance.total_chats      agent_performance.avg_response_time      agent_performance.avg_resolution_time      agent_performance.avg_rating      agent_perform
ance.Total_feedback
1      7/30/2022      Prerna Singh      11      0:00:38      0:04:20      4.11      9
2      7/30/2022      Nandani Gupta      11      0:01:15      0:28:25      3.14      7
3      7/30/2022      Aneya Jain      14      0:00:39      0:11:36      4.25      11
4      7/30/2022      Mahesh Sarade      14      0:01:04      0:15:46      4.71      7
5      7/30/2022      Swati      14      0:01:11      0:16:33      3.67      6
Time taken: 0.308 seconds, Fetched: 5 row(s)
hive> select count(date),agent from agent_logging_report group by agent limit 5;

```

Fig 1

3. List of all agents' names.

```

hive> select distinct agent_name from agent_logging_report;
Query ID = cloudera_20221008193030_f16c337f-5479-4d72-b1ce-bc141175650f
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_1665280805143_0001, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1665280805143_0001/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1665280805143_0001
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-08 19:30:30.017 Stage-1 map = 0%, reduce = 0%
2022-10-08 19:30:40.084 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.85 sec
2022-10-08 19:30:49.653 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 3.95 sec
MapReduce Total cumulative CPU time: 3 seconds 950 msec
Ended Job = job_1665280805143_0001
MapReduce Jobs Launched:
Stage-Stage1: Map: 1 Reduce: 1 Cumulative CPU: 3.95 sec HDFS Read: 63132 HDFS Write: 638 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 950 msec
OK
Aditya Shinde
Aditya Jot
Amersh
Aneya Jain
Ankitha
Anurag Tiwari
Aravind
Ayushi Mishra
Bharath
Baktiar Ahmed Bappy
Chaitra K Hiremath
Deeppranjan Gupta
Dibyanshu
Harikrishnan Shaji
Harishkesh Neogi
Hyder Abbas
Insuron Intelligence
Ishwant Kumar
Jawala Prakash
Jaydeep Dixit
Khushboo Priya
Madhulika G
Mahesh Sarade

```

Select distinct agent from agent_logging_report ;

4. Find out agent average rating.

```

hive> select Agent name,avg(avg_rating) from agent_performance group by agent name limit 5;
Query ID = cloudera_20221008203330_e074fec-34f8-420e-b080-2b5317becafe
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_1665280805143_0002, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1665280805143_0002/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1665280805143_0002
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-08 20:38:40.872 Stage-1 map = 0%, reduce = 0%
2022-10-08 20:39:04.758 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 5.39 sec
2022-10-08 20:39:20.617 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 9.24 sec
MapReduce Total cumulative CPU time: 9 seconds 240 msec
Ended Job = job_1665280805143_0002
MapReduce Jobs Launched:
Stage-Stage1: Map: 1 Reduce: 1 Cumulative CPU: 9.24 sec HDFS Read: 119549 HDFS Write: 102 SUCCESS
Total MapReduce CPU Time Spent: 9 seconds 240 msec
OK
Abhishek      0.0
Aditya      0.0
Aditya Shinde      1.8803333499627278
Aditya Jot      2.3453333377830135
Amersh      0.0
Time taken: 78.183 seconds, Fetched: 5 row(s)
hive>

```

select agent_name,avg(average_rating), from
agent_performance group by agent_name limit 5;

5. Total working days for each agents

```
Time taken: 42.964 seconds, Fetched: 5 row(s)
hive> select count(date) as count,date from agent_logging_report group by agent_name limit 5;
Query ID = cloudera_20221009191010_945ceb7-f57f-4a6b-a45a-daa42bfb920
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.reducers=<number>
Starting Job = job_1665366724303_0002, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1665366724303_0002/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1665366724303_0002
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-09 19:10:27.143 Stage:1 map = 0%, reduce = 0%
2022-10-09 19:10:37.788 Stage:1 map = 100%, reduce = 0%, Cumulative CPU 1.9 sec
2022-10-09 19:10:44.278 Stage:1 map = 100%, reduce = 100%, Cumulative CPU 3.94 sec
MapReduce Total cumulative CPU time: 3 seconds 940 msec
Ended Job = job_1665366724303_0002
MapReduce Jobs Launched:
Stage:Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 3.94 sec HDFS Read: 63988 HDFS Write: 238 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 940 msec
OK
count    agent
1        Aditya Shinde
4        Aditya Iot
4        Amersa
18       Aneya Jain
4        Anshika
37       Anurag Tiwari
18       Aravind
18       Arushi Mishra
9        Bharath
17       Baktar Ahmed Boppy
13       Chaitra K Hiremath
28       Deepranjan Gupta
388      Dilysandu
23       Harikrishnan Shaji
57       Hrisikesh Neogi
Time taken: 28.463 seconds, Fetched: 15 row(s)
hive>
```

Select count(date),agent from agent_logging_report group by
agent_name limit 5;

6. Total query that each agent have taken

```
Time taken: 51.964 seconds, Fetched: 1 row(s)
hive> select sum(total_chats) as sum,agent_name from agent_performance group by agent_name order by sum desc limit 5;
Query ID = cloudera_20221009211010_h025b07-bcfc-4663-8c09-5a1eefb7f
Total jobs = 2
Launching Job 1 out of 2
Number of reduce tasks not specified. Estimated from input data size: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reducers=<number>
Starting Job = job_1665280805143_0008, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1665280805143_0008/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1665280805143_0008
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-08 21:17:02.185 Stage:1 map = 0%, reduce = 0%
2022-10-08 21:17:14.729 Stage:1 map = 100%, reduce = 0%, Cumulative CPU 3.83 sec
2022-10-08 21:17:25.467 Stage:1 map = 100%, reduce = 100%, Cumulative CPU 8.56 sec
MapReduce Total cumulative CPU time: 8 seconds 560 msec
Ended Job = job_1665280805143_0008
Launching Job 2 out of 2
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reducers=<number>
Starting Job = job_1665280805143_0009, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1665280805143_0009/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1665280805143_0009
Hadoop job information for Stage-2: number of mappers: 1; number of reducers: 1
2022-10-08 21:17:50.947 Stage:2 map = 0%, reduce = 0%
2022-10-08 21:18:02.016 Stage:2 map = 100%, reduce = 0%, Cumulative CPU 3.18 sec
2022-10-08 21:18:17.798 Stage:2 map = 100%, reduce = 100%, Cumulative CPU 7.12 sec
MapReduce Total cumulative CPU time: 7 seconds 120 msec
Ended Job = job_1665280805143_0009
MapReduce Jobs Launched:
Stage:Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 8.56 sec HDFS Read: 118142 HDFS Write: 2388 SUCCESS
Stage:Stage-2: Map: 1 Reduce: 1 Cumulative CPU: 7.12 sec HDFS Read: 7288 HDFS Write: 74 SUCCESS
Total MapReduce CPU Time Spent: 15 seconds 680 msec
OK
sum    agent_name
578    Hrisikesh Neogi
460    Nandanji Gupta
542    Zeeshan
542    Hailtry
524    Smiti
Time taken: 93.714 seconds, Fetched: 5 row(s)
cloudera@quickstart~$
```

Select Sum(total_chats) as sum,agent_name
from agent_performance group by agent_name
Order by sum Desc limit 5;

7. Total Feedback that each agent have received

select sum(total_feedback) as sum,agent_name from
agent_performance group by agent_name

Order by sum desc limit 5;

```
Time taken: 0.196 seconds, Fetched: 5 row(s)
hive> select sum(total_feedback) as sum_agent_name from agent_performance group by agent_name order by sum desc limit 5;
Query ID = cloudera_20221008212222_af1e7134-6f06-47fa-a34f-ct3bc7977027
Total jobs = 2
Launching Job 1 out of 2
Number of reduce tasks not specified. Estimated from input data size: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reducers=<number>
Starting Job = job_1665280805143_0010, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1665280805143_0010/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1665280805143_0010
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-08 21:22:22,496 Stage-1 map = 0%, reduce = 0%
2022-10-08 21:22:35,892 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 4.52 sec
2022-10-08 21:22:52,812 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 8.73 sec
MapReduce Total cumulative CPU time: 8 seconds 730 msec
Ended Job = job_1665280805143_0010
Launching Job 2 out of 2
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reducers=<number>
Starting Job = job_1665280805143_0011, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1665280805143_0011/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1665280805143_0011
Hadoop job information for Stage-2: number of mappers: 1; number of reducers: 1
2022-10-08 21:23:09,887 Stage-2 map = 0%, reduce = 0%
2022-10-08 21:23:21,129 Stage-2 map = 100%, reduce = 0%, Cumulative CPU 3.09 sec
2022-10-08 21:23:39,313 Stage-2 map = 100%, reduce = 100%, Cumulative CPU 7.67 sec
MapReduce Total cumulative CPU time: 7 seconds 670 msec
Ended Job = job_1665280805143_0011
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 8.73 sec HDFS Read: 118234 HDFS Write: 2295 SUCCESS
Stage-Stage-2: Map: 1 Reduce: 1 Cumulative CPU: 7.67 sec HDFS Read: 7187 HDFS Write: 76 SUCCESS
Total MapReduce CPU Time Spent: 16 seconds 400 msec
OK
sum agent_name
367 Hrisikesh Neogi
364 Mithun S
347 Maitry
335 Zeeshan
329 Ayushi Mishra
Time taken: 94.924 seconds, Fetched: 5 row(s)
```

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

```
Time taken: 94.924 seconds, Fetched: 5 row(s)
hive> select agent_name,avg_rating from agent_performance where avg_rating between 3.5 and 4;
OK
agent_name avg_rating
Manjunatha A 3.6
Boktiar Ahmed Bappy 4.0
Prateek Jot 3.75
Nandani Gupta 3.79
Jaydeep Dixit 3.95
Himesh Sarade 3.84
Zeeshan 3.79
Hrisikesh Neogi 3.77
Muskan Garg 4.0
Khushboo Priya 3.79
Mezlin 3.95
Jawala Prakash 3.89
Shiva Srivastava 4.0
Nishtha Jain 3.67
Maitry 4.0
Shiva Srivastava 3.61
Prerna Singh 3.8
Shivan K 4.0
Boktiar Ahmed Bappy 3.8
Muskan Garg 3.62
Chaitra K Hiremath 4.0
Harikrishnan Shaji 3.73
Mezlin 3.92
Deepranjan Gupta 3.69
Himesh Sarade 3.81
Nishtha Jain 3.92
Shivan K 3.9
Madhulika G 3.62
Rishav Dash 3.8
Prabir Kumar Satapathy 4.0
Boktiar Ahmed Bappy 4.0
Maitry 3.73
Nishtha Jain 3.69
Khushboo Priya 3.95
Boktiar Ahmed Bappy 3.77
Himesh Sarade 3.5
Rishav Dash 3.67
Mithun S 3.75
Jawala Prakash 3.68
Aditya Jot 3.86
Swati 3.77
Nandani Gupta 3.73
```

Select agent_name from agent_performance where average_rating BETWEEN 3.5 AND 4;

9. Agent name who have rating less than 3.5 .

Select agent_name from agent_performance where average_rating < 3.5 limit 10;

```
Time taken: 0.457 seconds, Fetched: 114 row(s)
hive> select agent_name,avg_rating from agent_performance where avg_rating < 3.5 limit 10;
OK
agent_name avg_rating
Nandani Gupta 3.14
Hitesh Choudhary 0.0
Saijeon 0.0
Anirudh 0.0
Shiva Srivastava 0.0
Nityesh 0.0
Akshish 0.0
Ajay Mishra 0.0
Aditya Shinde 0.0
Jayant Kumar 0.0
Time taken: 0.493 seconds, Fetched: 10 row(s)
```

10. Agent name who have rating more than 4.5.

Select agent_name from agent_performance where average_rating > 4.5 limit 5;

```
Time taken: 0.493 seconds, Fetched: 10 row(s)
hive> select agent_name,avg_rating from agent_performance where avg_rating > 4.5 limit 10;
OK
agent_name      avg_rating
-----
Ajaya Jain      4.75
Rohesh Sarade  4.71
Rohesh          4.62
Srikumarreddy N 5.0
Sanjeev Kumar   5.0
Nagjyothan Shaji 4.57
Sumanliya Silvakumar 4.75
Rakti Arvind Bapay 4.75
Shivarama Srinivas 5.0
Likhawant Kumar 4.67
Time taken: 0.381 seconds, Fetched: 10 row(s)
hive>
```

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

```
11.186666666666667 zreshan
Time taken: 49.584 seconds, Fetched: 31 row(s)
hive> select count(*) as feedback from agent_performance group by agent_name having average>4.5 limit 5;
OK
count(*)
-----
1
Time taken: 0.001 seconds, Fetched: 1 row(s)
hive>
Query ID = cloudera_20221008215757_10257072-5078-4493-96c8-e3ad307de873
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.reducers=number
Starting job = job_16652880805143_0016, Tracking URL = http://quickstart.cloudera:8088/proxy/application_16652880805143_0016/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_16652880805143_0016
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-08 21:57:31,119 Stage-1 map = 0%, reduce = 0%
2022-10-08 21:57:44,727 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 4.67 sec
2022-10-08 21:58:00,255 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 9.97 sec
MapReduce Total cumulative CPU time: 9 seconds 970 msec
Ended Job = job_16652880805143_0016
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 9.97 sec HDFS Read: 120156 HDFS Write: 120 SUCCESS
Total MapReduce CPU Time Spent: 9 seconds 970 msec
OK
average agent_name
-----
5.1 Aditya Shinde
7.6 Ajaya Jain
7.666666666666667 Aravind
10.966666666666667 Ayushi Mishra
0.7333333333333333 Bharrath
Time taken: 46.443 seconds, Fetched: 5 row(s)
hive>
```

Select avg(total_feedback) as average, agent_name from agent_performance group by agent_name having average>4.5;

12. Average weekly response time for each agent.

Here, to find average weekly response time, we have to first convert the “avg_response_time” column into seconds and then convert it into week. Since the column is in the format hh:mm:ss, I have first converted it into seconds and then divided it by (60*60*24*7) to convert it into week format.

Select a.agent_name,
avg(col1[0]*3600+col1[1]*60+col1[2])/604800 from (select
Agent_name,split(avg_response_time,':') as col1 from
agent_performance) a
Group by a.agent_name;

```

Time taken: 28.463 seconds, Fetched: 10 row(s)
hive> SELECT a.agent_name,avg(col1[0]*3600+col1[1]*60+col1[2])/604800 from (select agent_name,split(avg_response_time,':') as col1 from agent_performance)a group by a.agent_name;
Query ID = cloudera_20221009193839_cfee8fde-e532-4674-9f1a-a13c68762f7
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_1665366724383_0803, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1665366724383_0803/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1665366724383_0803
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-09 19:38:28.922 Stage-1 map = 0%, reduce = 0%
2022-10-09 19:38:39.770 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.6 sec
2022-10-09 19:38:49.307 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 5.16 sec
MapReduce Total cumulative CPU time: 5 seconds 168 msec
Ended Job = job_1665366724383_0803
MapReduce Jobs Launched:
Stage:Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 5.16 sec HDFS Read: 122082 HDFS Write: 2046 SUCCESS
Total MapReduce CPU Time Spent: 5 seconds 168 msec
OK
a.agent_name      c1
Abhishek          0.0
Aditya            0.0
Aditya Shinde     4.916225749559083E-5
Aditya Iot        5.616181657848325E-5
Anursh            0.0
Anaya Jain        3.4942686776014103E-5
Anirudh          3.598985896052557E-5
Ankit Sharma      0.0
Ankitha           7.330246913586247E-6
Anurag Tiwari     1.384480327336806E-5
Aravind           3.5326483245149913E-5
Ashad Nasim       6.38778059611993E-5
Ashish            0.0
Ayush Mishra     9.975749555862893E-5
Bharath           4.43112693226931E-5
Baktiar Ahmed Bappy 1.091820987654321E-4
Chaitra K Hiremath 2.562245055379108E-5
Dhanraj Gupta     8.78029262829297E-5
Dibyanshu         2.0943562616229278E-6
Harikrishnan Shaji 5.616181657848325E-5
Hitesh Choudhary  0.0
Hrishikesh Negi   8.349867724867724E-5
Iyfer Abbas       0.0

```

13. Average weekly resolution time for each agents .

Here also,we have to do the same thing as in previous question, the only difference is here we have to use “avg_resolution-time”.

Select a.agent_name,
 $\text{avg}(\text{col1}[0]*3600+\text{col1}[1]*60+\text{col1}[2])/604800$ from (select
 Agent_name,split(avg_resolution_time,':') as col1 from
 agent_performance) a
 Group by a.agent_name limit 10;

```

Time taken: 33.344 seconds, Fetched: 70 row(s)
hive> SELECT a.agent_name,avg(col1[0]*3600+col1[1]*60+col1[2])/604800 from (select agent_name,split(avg_resolution_time,':') as col1 from agent_performance)a group by a.agent_name limit 10;
Query ID = cloudera_20221009195151_87b791ea-9145-455f-8a26-b9343a5be9f9
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_1665366724383_0804, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1665366724383_0804/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1665366724383_0804
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-09 19:51:27.709 Stage-1 map = 0%, reduce = 0%
2022-10-09 19:51:29.613 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.89 sec
2022-10-09 19:51:48.390 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 5.42 sec
MapReduce Total cumulative CPU time: 5 seconds 420 msec
Ended Job = job_1665366724383_0804
MapReduce Jobs Launched:
Stage:Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 5.42 sec HDFS Read: 122295 HDFS Write: 258 SUCCESS
Total MapReduce CPU Time Spent: 5 seconds 420 msec
OK
a.agent_name      c1
Abhishek          0.0
Aditya            0.0
Aditya Shinde     0.081026179453262786E
Aditya Iot        9.743716931216931E-4
Anursh            0.0
Anaya Jain        5.43154761984762E-4
Anirudh          3.06657848324515E-4
Ankit Sharma      0.0
Ankitha           9.085731922398589E-5
Anurag Tiwari     1.22242680776014E-4
Time taken: 33.576 seconds, Fetched: 10 row(s)
hive>

```


14. Find the number of chat on which they have received a feedback .

```

hive> select agent_name,sum(total_chats),total_feedback
from agent_performance where total_feedback group by agent_name,total_feedback limit 10;
Query ID = cloudera-2022100822222_89c289c-ed1e-48ca-97e1-ec930e8fb28
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.reducers=number
Starting Job = job_1665280805143_0019, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1665280805143_0019/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1665280805143_0019
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-08 22:22:49,751 Stage-1 map = 0%, reduce = 0%, Cumulative CPU 5.36 sec
2022-10-08 22:23:04,184 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 9.63 sec
MapReduce Total cumulative CPU time: 9 seconds 630 msec
Ended Job = job_1665280805143_0019
MapReduce Jobs Launched:
Stage:Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 9.63 sec HDFS Read: 119789 HDFS Write: 190 SUCCESS
Total MapReduce CPU Time Spent: 9 seconds 630 msec
OK
agent_name      sum      total_feedback
Aditya Shinde   8         7
Aditya Shinde   17         8
Aditya Shinde   67         9
Aditya Shinde   18        11
Aditya Shinde   27        14
Aditya Shinde   49        15
Aditya Shinde   34        19
Aditya Shinde   57        28
Aditya Iot      11         2
Aditya Iot      18         3
Time taken: 56.119 seconds, Fetched: 10 row(s)
hive>

```

Select agent_name,sum(total_chats),total_feedback
from agent_performance where total_feedback>0
group by agent_name,total_feedback limit 10;

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

Select a.agent,
avg(col1[0]*3600+col1[1]*60+col1[2])/3600 from (select
agent,split(duration,':') as col1 from agent_performance) a
Group by a.agent_name limit 10;

```

hive> select a.agent_name,sum(col1[0]*3600+col1[1]*60+col1[2])/3600 from (select agent_name,split(duration,':') as col1 from agent_performance) a group by a.agent_name limit 10;
FAILED: SemanticException [Error 10004]: line 1:75 Invalid table alias or column reference 'agent_name': (possible column names are: sno, agent, date, login time, logout time, duration)
hive> select a.agent_name,sum(col1[0]*3600+col1[1]*60+col1[2])/3600 from (select agent,split(duration,':') as col1 from agent_performance) a group by a.agent_name limit 10;
Query ID = cloudera-20221009195859_40655af4-16ea-472f-ad86-32eb29aebf7
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.reducers=number
Starting Job = job_1665366724303_0005, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1665366724303_0005/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1665366724303_0005
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-09 19:58:22,210 Stage-1 map = 0%, reduce = 0%, Cumulative CPU 2.47 sec
2022-10-09 19:58:37,826 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 5.01 sec
MapReduce Total cumulative CPU time: 5 seconds 10 msec
Ended Job = job_1665366724303_0005
MapReduce Jobs Launched:
Stage:Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 5.01 sec HDFS Read: 66860 HDFS Write: 301 SUCCESS
Total MapReduce CPU Time Spent: 5 seconds 10 msec
OK
a.agent      c1
Aditya Shinde   0.03611111111111111
Aditya Iot      15.731111111111112
Anurag Tiwari   3.063888888888889
Ananya Jain     42.875555555555555
Aditya         2.265944444444444
Anurag Tiwari   2.778888888888889
Aravind 24.299166666666666
Ayushi Mishra   10.121666666666667
Bharath 48.876666666666667
Nattier Ahmed Bhojwani 40.26861111111111
Time taken: 29.33 seconds, Fetched: 10 row(s)
hive>

```

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> select alr.agent,ap.total_chats,ap.avg_rating from hive_assignment.agent_logging_report alr
> inner join hive_assignment.agent_performance ap
> on alr.agent=ap.agent_name limit 5;
Query ID = cloudera.20221008223838_e6733017-e90e-4c7b-8f22-2de767519b18
Total jobs = 1
Execution log at: /tmp/cloudera/cloudera.20221008223838_e6733017-e90e-4c7b-8f22-2de767519b18.log
2022-10-08 10:38:25 Starting to launch local task to process map join; maximum memory = 932184064
2022-10-08 10:38:30 Dump the side-table for tag: 0 with group count: 49 into file: file:/tmp/cloudera/c487f6b7-77ae-4996-b590-971659c195ef/hive_2022-10-08_22-38-14_575_3209597309293893201-1/-local-10003/HashTable-Stage-3/MapJoin-mapfile000-..hashtable (5560 bytes)
2022-10-08 10:38:30 End of local task; Time Taken: 4.568 sec.
Execution completed successfully
HadoopLocal 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_1665288805143_0020, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1665288805143_0020/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1665288805143_0020
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0
2022-10-08 22:39:05.579 Stage-3 map = 0%, reduce = 0%, Cumulative CPU 6.03 sec
MapReduce Total cumulative CPU time: 6 seconds 30 msec
Ended Job = job_1665288805143_0020
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1 Cumulative CPU: 6.03 sec HDFS Read: 11797 HDFS Write: 115 SUCCESS
Total MapReduce CPU Time Spent: 6 seconds 30 msec
OK
alr.agent ap.total_chats ap.total_feedback ap.avg_rating
Prerna Singh 11 9 4.11
Prerna Singh 11 9 4.11
Prerna Singh 11 9 4.11
Prerna Singh 11 9 4.11
Prerna Singh 11 9 4.11
Time taken: 52.304 seconds, Fetched: 5 row(s)
hive>
```

Select alr.agent,ap.total_chats,ap.avg_rating,ap.total_feedback
From hive_assignment.agent_logging_report alr
Inner join hive_assignment.agent_performance ap
On alr.agent=ap.agent_name limit 5;

● LEFT JOIN-

```
hive> select alr.agent,ap.total_chats,ap.avg_rating,ap.total_feedback from hive_assignment.agent_logging_report alr
left join hive_assignment.agent_performance ap
on alr.agent=ap.agent_name limit 5;
Query ID = cloudera.20221008224040_ef601714-26cc-46b9-a96c-7ec917f6786f
Total jobs = 1
Execution log at: /tmp/cloudera/cloudera.20221008224040_ef601714-26cc-46b9-a96c-7ec917f6786f.log
2022-10-08 10:41:10 Starting to launch local task to process map join; maximum memory = 932184064
2022-10-08 10:41:14 Dump the side-table for tag: 1 with group count: 70 into file: file:/tmp/cloudera/c487f6b7-77ae-4996-b590-971659c195ef/hive_2022-10-08_22-40-57_092_4930451372819752082-1/-local-10003/HashTable-Stage-3/MapJoin-mapfile11-..hashtable (25920 bytes)
2022-10-08 10:41:14 End of local task; Time Taken: 3.68 sec.
Execution completed successfully
HadoopLocal 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_1665288805143_0021, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1665288805143_0021/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1665288805143_0021
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0
2022-10-08 22:41:32.868 Stage-3 map = 0%, reduce = 0%
2022-10-08 22:41:47.564 Stage-3 map = 100%, reduce = 0%, Cumulative CPU 4.8 sec
MapReduce Total cumulative CPU time: 4 seconds 800 msec
Ended Job = job_1665288805143_0021
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1 Cumulative CPU: 4.8 sec HDFS Read: 11767 HDFS Write: 142 SUCCESS
Total MapReduce CPU Time Spent: 4 seconds 800 msec
OK
alr.agent ap.total_chats ap.total_feedback ap.avg_rating
Shivananda Somwane 4 1 5.0
Shivananda Somwane 14 9 4.67
Shivananda Somwane 5 4 5.0
Shivananda Somwane 20 10 4.22
Shivananda Somwane 24 14 5.0
Time taken: 33.038 seconds, Fetched: 5 row(s)
hive>
```

Select alr.agent,ap.total_chats,ap.avg_rating,ap.total_feedback
From hive_assignment.agent_logging_report alr
left join hive_assignment.agent_performance ap
On alr.agent=ap.agent_name limit 5;

● RIGHT JOIN-

```

Shivananda Sonwane 24 14 5.0
Time taken: 53.029 seconds, Fetched: 3 row(s)
Hive> select alr.agent,ap.total_chats,ap.total_feedback,ap.avg_rating from hive_assignment.agent_logging_report alr
> right join hive_assignment.agent_performance ap
> on alr.agent=ap.agent_name limit 5;
Query ID = cloudera.20221008224242.c9313773-8a55-4b43-b686-ff8706f38d6
Total jobs = 1
Execution log at: /tmp/cloudera/cloudera.20221008224242.c9313773-8a55-4b43-b686-ff8706f38d6.log
2022-10-08 10:43:00 Starting to launch local task to process map join; maximum memory = 932184064
2022-10-08 10:43:03 Dump the side-table for tag: 0 with group count: 49 into file: file:/tmp/cloudera/c487f6b7-77ae-4996-b599-971659c195ef/hive_2022-10-08_22-42-48_152_246980289597614783-1/-local_10003/HashTable-Stage-3/MapJoin-map
j1a29--hashtable
2022-10-08 10:43:03 Uploaded file to: file:/tmp/cloudera/c487f6b7-77ae-4996-b599-971659c195ef/hive_2022-10-08_22-42-48_152_246980289597614783-1/-local_10003/HashTable-Stage-3/MapJoin-mapfilec20--hashtable (5560 bytes)
2022-10-08 10:43:03 End of local task, Time Taken: 3.295 sec.
Execution completed successfully
MapReduce 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_1665288805143_0022, Tracking URL = http://quickstart.cloudera:8080/proxy/application_1665288805143_0022/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1665288805143_0022
Hadoop job information for Stage:3: number of mappers: 1; number of reducers: 0
2022-10-08 22:43:29.188 Stage:3 map = 0%, reduce = 0%
2022-10-08 22:43:37.155 Stage:3 map = 100%, reduce = 0%, Cumulative CPU 4.82 sec
MapReduce Total cumulative CPU time: 4 seconds 820 msec
Ended Job = job_1665288805143_0022
MapReduce Jobs Launched:
Stage:Stage=3: Map: 1, Cumulative CPU: 4.82 sec, HDFS Read: 11768 HDFS Write: 115 SUCCESS
Total MapReduce CPU Time Spent: 4 seconds 820 msec
OK
alr.agent ap.total_chats ap.total_feedback ap.avg_rating
Prerna Singh 11 9 4.11
Prerna Singh 11 9 4.11
Prerna Singh 11 9 4.11
Prerna Singh 11 9 4.11
Prerna Singh 11 9 4.11
Time taken: 50.165 seconds, Fetched: 5 row(s)
hive>

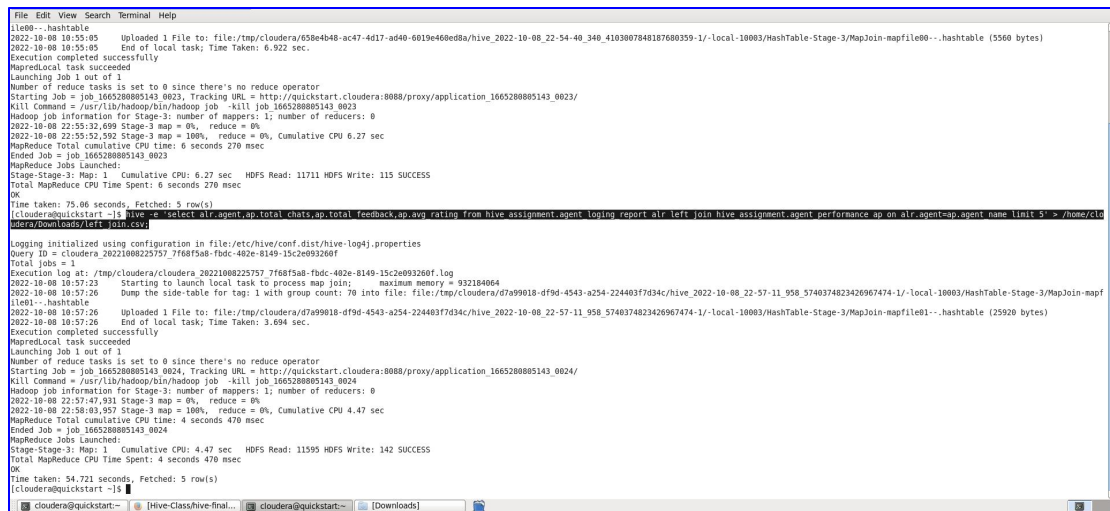
```

Select alr.agent,ap.total_chats,ap.avg_rating,ap.total_feedback
 From hive_assignment.agent_logging_report alr
 right join hive_assignment.agent_performance ap
 On alr.agent=ap.agent_name limit 5;

Now, we have to export the joined output in the local HDFS system. For this, we have to exit the hive terminal or start a new terminal and then perform the following query.

hive -e 'Select alr.agent,alr.date,ap.total_chats,ap.total_feedback
 from hive_assignment2.agent_logging_report alr join
 hive_assignment2.agent_performance ap on alr.agent =
 ap.agent_name limit 5' >
 /home/cloudera/hive_Assignment2/inner.join.csv;

hive -e 'Select alr.agent,alr.date,ap.total_chats,ap.total_feedback
 from hive_assignment2.agent_logging_report alr left join
 hive_assignment2.agent_performance ap on alr.agent =
 ap.agent_name limit 5' >
 /home/cloudera/hive_Assignment2/left.join.csv;



```
hive -e 'Select alr.agent,alr.date,ap.total_chats,ap.total_feedback
from hive_assignment2.agent_logging_report alr right join
hive_assignment2.agent_performance ap on alr.agent =
ap.agent_name limit 5' >
/home/cloudera/hive_Assignment2/right.join.csv;
```

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

First of all, I have performed partitioning and Bucketing for Agent_logging_report table.

```
Create table alr_part_bucket(
Sno int, Date string,login_time string,logout_time string,
Duration string) partitioned by (agent string)
Clustered by (Date) sorted by (Date) into 4 buckets
Row format delimited
Fields terminated by ',';
```

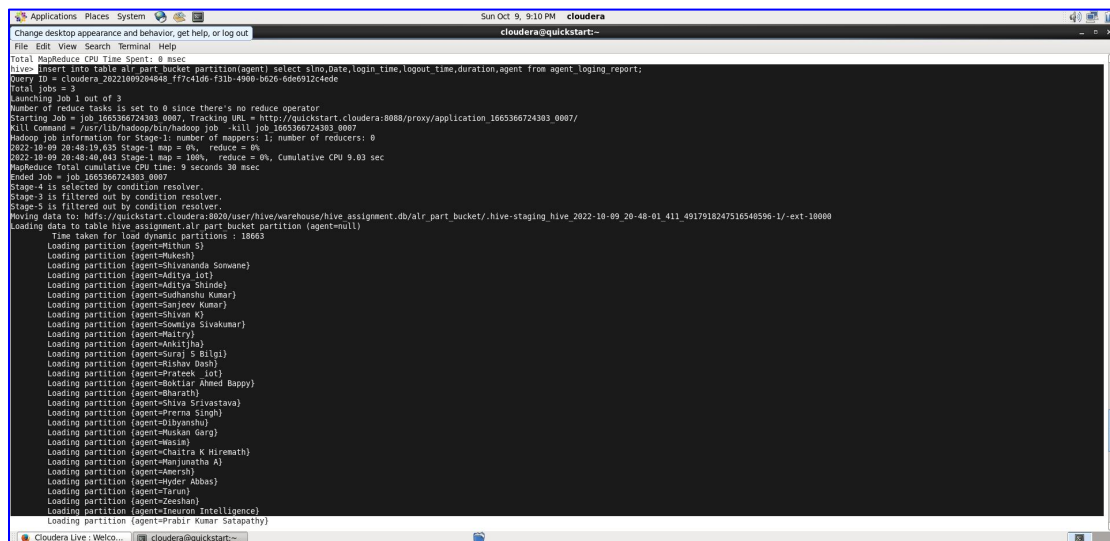
Then set the Hive Dynamic mode partition.

```
Set hive.exec.dynamic.partition=true;
Set hive.exec.dynamic.partition.mode=nonstrict;
```

And then loading the data into the table using insert overwrite command.

```
doctlar Ahmed Bappy 40.26861111111111
Time taken: 29.33 seconds, Fetched: 10 rows)
hive> create table alr_part_bucket(
> sln int, Date string, login_time string,
> logout_time string, duration string,
> partitioned by (agent string)
> clustered by (Date) sorted by (Date) into 4 Buckets
> row format delimited
> fields terminated by ',';
OK
Time taken: 1.202 seconds
hive> set hive.exec.dynamic.partition=true;
hive> set hive.exec.dynamic.partition.mode=strict;
hive> insert into table alr_part_bucket partition(agent) select slno,Agent,Date,login_time,logout_time,duration from agent_logging_report;
Query ID = cloudera_20221009202929_6e09e274-58c3-4756-bc49-080a8a267f33
```

Insert into table alr_part_bucket partition (agent)
Select slno,Date,login_time,logout_time,duration,agent from
agent_logging_report;



From the above image, it can be seen that the partitioning and bucketing part is created successfully.
Now , the steps have to be repeated for the second table
“Agent_performance”.

Create table AP_partition_Bucket(
sln int,Date date,
Total_chat string,
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 8 BUCKETS
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ',';

Insert into table ap_part_bucket partition (agent_name)
Select
sno,date,total_chats,avg_response_time,avg_resolution_time,
Avg rating,total feedback,Agent_name
from
agent_performance;

```
Applications Places System Sun Oct 9, 9:24 PM cloudera
Change desktop appearance and behavior, get help, or log out
cloudera@quickstart:~$
File Edit View Search Terminal Help
hive> Create table ap_part_bucket
> (sno int, date string, total_chats int,
> avg_response_time int, avg_resolution_time int,
> avg_rating float, total_feedback int)
> partitioned by (agent_name string)
> clustered by (date) sorted by (date) into 8 buckets
> row format delimited
> fields terminated by ',';
OK
Time taken: 0.176 seconds
hive> insert into table ap_part_bucket partition (agent) select sno,date,total_chats,avg_response_time,avg_resolution_time,avg_rating,total_feedback,Agent_name from agent_performance;
FAILED: SemanticException Partition spec (agent=null) contains non-partition columns
hive> insert into table ap_part_bucket partition (agent_name) select sno,date,total_chats,avg_response_time,avg_resolution_time,avg_rating,total_feedback,Agent_name from agent_performance;
Query ID = cloudera-20221009212121.940a2906-c8fe-44f7-a288-8360c722786
Total jobs = 3
Launching Job 1 out of 3
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1665366724383_0008, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1665366724383_0008/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1665366724383_0008
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 0
2022-10-09 21:22:24.849 Stage:1 map = 0%, reduce = 0%
2022-10-09 21:22:53.955 Stage:1 map = 100%, reduce = 0%, Cumulative CPU 10.36 sec
MapReduce Total cumulative CPU time: 10 seconds 768 msec
Ended Job = job_1665366724383_0008
Stage-4 is selected by condition resolver.
Stage-3 is filtered out by condition resolver.
Stage-5 is filtered out by condition resolver.
Moving data to: hdfs://quickstart.cloudera:8020/user/hive/warehouse/hive_assignment.db/ap_part_bucket/.hive-staging_hive_2022-10-09_21-21-59_046_7456821624314341772-1/-ext-10000
Loading data to table hive_assignment.ap_part_bucket partition (agent_name=null)
Time taken for load dynamic partitions : 19563
Loading partition (agent_name=Himanshu K)
Loading partition (agent_name=Mahak )
Loading partition (agent_name=Shivansha Somwane)
Loading partition (agent_name=Aditya Jot )
Loading partition (agent_name=Maity )
Loading partition (agent_name=Prateek Kumar)
Loading partition (agent_name=Prateek Jot )
Loading partition (agent_name=Hakim Ahmed Bappy)
Loading partition (agent_name=Handant Gupta)
Loading partition (agent_name=Uday Mishra)
Loading partition (agent_name=Mithun S)
Loading partition (agent_name=Hyder Abbas)
Loading partition (agent_name=Mukesh )
Loading partition (agent_name=Smriti Srivastava)
Loading partition (agent_name=Neeraj Intelligence )
Loading partition (agent_name=Amresh )
Loading partition (agent_name=Deeptaran Gupta)
Loading partition (agent_name=Harikrishnan Shaji)
```

```
Applications Places System Sun Oct 9, 9:24 PM cloudera
Browse and run installed applications
cloudera@quickstart:~$
File Edit View Search Terminal Help
Time taken for adding to write entity : 98
Partition hive_assignment.ap_part_bucket(partition=Abhishek ) stats: [numFiles=1, numRows=30, totalSize=846, rawDataSize=816]
Partition hive_assignment.ap_part_bucket(partition=Aditya ) stats: [numFiles=1, numRows=30, totalSize=846, rawDataSize=816]
Partition hive_assignment.ap_part_bucket(partition=Aditya Shinde) stats: [numFiles=1, numRows=30, totalSize=874, rawDataSize=844]
Partition hive_assignment.ap_part_bucket(partition=Aditya Jot ) stats: [numFiles=1, numRows=30, totalSize=874, rawDataSize=844]
Partition hive_assignment.ap_part_bucket(partition=Amresh ) stats: [numFiles=1, numRows=30, totalSize=846, rawDataSize=816]
Partition hive_assignment.ap_part_bucket(partition=Ananya Jain) stats: [numFiles=1, numRows=30, totalSize=880, rawDataSize=850]
Partition hive_assignment.ap_part_bucket(partition=Anirudh ) stats: [numFiles=1, numRows=30, totalSize=854, rawDataSize=824]
Partition hive_assignment.ap_part_bucket(partition=Ankit Sharma) stats: [numFiles=1, numRows=30, totalSize=846, rawDataSize=816]
Partition hive_assignment.ap_part_bucket(partition=Anshika ) stats: [numFiles=1, numRows=30, totalSize=846, rawDataSize=816]
Partition hive_assignment.ap_part_bucket(partition=Anurag Tiwari) stats: [numFiles=1, numRows=30, totalSize=846, rawDataSize=816]
Partition hive_assignment.ap_part_bucket(partition=Arvind ) stats: [numFiles=1, numRows=30, totalSize=844, rawDataSize=814]
Partition hive_assignment.ap_part_bucket(partition=Ashad Nasin) stats: [numFiles=1, numRows=30, totalSize=847, rawDataSize=817]
Partition hive_assignment.ap_part_bucket(partition=Ashish ) stats: [numFiles=1, numRows=30, totalSize=892, rawDataSize=862]
Partition hive_assignment.ap_part_bucket(partition=Ayushi Mishra) stats: [numFiles=1, numRows=30, totalSize=894, rawDataSize=864]
Partition hive_assignment.ap_part_bucket(partition=Bharath ) stats: [numFiles=1, numRows=30, totalSize=894, rawDataSize=864]
Partition hive_assignment.ap_part_bucket(partition=Harishesh Neogi) stats: [numFiles=1, numRows=30, totalSize=891, rawDataSize=861]
Partition hive_assignment.ap_part_bucket(partition=Chaitra K Hiremath) stats: [numFiles=1, numRows=30, totalSize=854, rawDataSize=824]
Partition hive_assignment.ap_part_bucket(partition=Deeptaran Gupta) stats: [numFiles=1, numRows=30, totalSize=895, rawDataSize=865]
Partition hive_assignment.ap_part_bucket(partition=Deeptaran Gupta) stats: [numFiles=1, numRows=30, totalSize=846, rawDataSize=816]
Partition hive_assignment.ap_part_bucket(partition=Deeptaran Gupta) stats: [numFiles=1, numRows=30, totalSize=888, rawDataSize=858]
Partition hive_assignment.ap_part_bucket(partition=Harikrishnan Shaji) stats: [numFiles=1, numRows=30, totalSize=888, rawDataSize=858]
Partition hive_assignment.ap_part_bucket(partition=Hitesh Choudhary) stats: [numFiles=1, numRows=30, totalSize=845, rawDataSize=815]
Partition hive_assignment.ap_part_bucket(partition=Hitesh Singh) stats: [numFiles=1, numRows=30, totalSize=862, rawDataSize=832]
Partition hive_assignment.ap_part_bucket(partition=Hyder Abbas) stats: [numFiles=1, numRows=30, totalSize=846, rawDataSize=816]
Partition hive_assignment.ap_part_bucket(partition=Divyanshu ) stats: [numFiles=1, numRows=30, totalSize=844, rawDataSize=814]
Partition hive_assignment.ap_part_bucket(partition=Ishwari Kumar) stats: [numFiles=1, numRows=30, totalSize=886, rawDataSize=856]
Partition hive_assignment.ap_part_bucket(partition=Jawala Prakash) stats: [numFiles=1, numRows=30, totalSize=896, rawDataSize=866]
Partition hive_assignment.ap_part_bucket(partition=Jayant Kumar) stats: [numFiles=1, numRows=30, totalSize=880, rawDataSize=850]
Partition hive_assignment.ap_part_bucket(partition=Jaydeep Dixit) stats: [numFiles=1, numRows=30, totalSize=903, rawDataSize=873]
Partition hive_assignment.ap_part_bucket(partition=Khushboo Priya) stats: [numFiles=1, numRows=30, totalSize=900, rawDataSize=870]
Partition hive_assignment.ap_part_bucket(partition=Madhulika G) stats: [numFiles=1, numRows=30, totalSize=901, rawDataSize=871]
Partition hive_assignment.ap_part_bucket(partition=Madhulika G) stats: [numFiles=1, numRows=30, totalSize=844, rawDataSize=814]
Partition hive_assignment.ap_part_bucket(partition=Maity ) stats: [numFiles=1, numRows=30, totalSize=890, rawDataSize=860]
Partition hive_assignment.ap_part_bucket(partition=Maity ) stats: [numFiles=1, numRows=30, totalSize=890, rawDataSize=860]
Partition hive_assignment.ap_part_bucket(partition=Mayunatha A) stats: [numFiles=1, numRows=30, totalSize=890, rawDataSize=860]
Partition hive_assignment.ap_part_bucket(partition=Mithun S) stats: [numFiles=1, numRows=30, totalSize=888, rawDataSize=858]
Partition hive_assignment.ap_part_bucket(partition=Mithun S) stats: [numFiles=1, numRows=30, totalSize=847, rawDataSize=817]
Partition hive_assignment.ap_part_bucket(partition=Mukesh Rao ) stats: [numFiles=1, numRows=30, totalSize=852, rawDataSize=822]
Partition hive_assignment.ap_part_bucket(partition=Muskan Gargi) stats: [numFiles=1, numRows=30, totalSize=902, rawDataSize=872]
Partition hive_assignment.ap_part_bucket(partition=Handant Gupta) stats: [numFiles=1, numRows=30, totalSize=892, rawDataSize=862]
Partition hive_assignment.ap_part_bucket(partition=Mishra Jain) stats: [numFiles=1, numRows=30, totalSize=888, rawDataSize=858]
Partition hive_assignment.ap_part_bucket(partition=Prerna Singh) stats: [numFiles=1, numRows=30, totalSize=889, rawDataSize=859]
Partition hive_assignment.ap_part_bucket(partition=Mishav Dash) stats: [numFiles=1, numRows=60, totalSize=1741, rawDataSize=1681]
Partition hive_assignment.ap_part_bucket(partition=Rohan ) stats: [numFiles=1, numRows=30, totalSize=846, rawDataSize=816]
```