

1. Create a schema based on the given dataset

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

```
[cloudera@quickstart:~]$ cd /tmp/hive_class_local_tmp
```

```
[cloudera@quickstart ~]$ cd /tmp/hive_class_local_tmp  
[cloudera@quickstart hive_class_local_tmp]$ ls  
AgentLoggingReport.csv AgentPerformance.csv CustomersCSV.csv dept_csv.txt sales_order_data_csv.txt  
[cloudera@quickstart hive_class_local_tmp]$
```

```
[cloudera@quickstart:~]
```

```
[cloudera@quickstart ~]$ cd /tmp/hive_class_local_tmp  
[cloudera@quickstart hive_class_local_tmp]$ ls  
AgentLoggingReport.csv AgentPerformance.csv CustomersCSV.csv dept_csv.txt sales_order_data_csv.txt  
[cloudera@quickstart hive_class_local_tmp]$ cd  
[cloudera@quickstart ~]$ hdfs dfs -put /tmp/hive_class_local_tmp/AgentLoggingReport.csv /  
[cloudera@quickstart ~]$ hdfs dfs -put /tmp/hive_class_local_tmp/AgentPerformance.csv /  
[cloudera@quickstart ~]$ hdfs dfs -ls /
```

Found 16 items

```
-rw-r--r-- 1 cloudera supergroup 0 2022-08-29 04:12 /.autofsck  
-rw-r--r-- 1 cloudera supergroup 0 2022-08-29 04:12 /.autorelabel  
-rw-r--r-- 1 cloudera supergroup 53911 2023-01-03 01:15 /AgentLoggingReport.csv  
-rw-r--r-- 1 cloudera supergroup 112661 2023-01-03 01:16 /AgentPerformance.csv  
---xr--r-- 1 cloudera supergroup 59 2022-08-29 04:51 /LocalFile  
drwxrwxrwx - hdfs supergroup 0 2017-10-23 09:15 /benchmarks  
drwxr-xr-x - cloudera supergroup 0 2022-08-29 04:12 /bin  
drwxr-xr-x - cloudera supergroup 0 2022-08-29 04:12 /boot  
drwxr-xr-x - cloudera supergroup 0 2022-08-29 04:12 /dev  
drwxr-xr-x - hbase supergroup 0 2023-01-03 01:05 /hbase  
drwxr-xr-x - cloudera supergroup 0 2022-09-10 09:54 /home  
drwxr-xr-x - solr solr 0 2017-10-23 09:18 /solr  
-rw-r--r-- 1 cloudera supergroup 26 2022-12-27 02:41 /test.txt  
drwxrwxrwt - hdfs supergroup 0 2022-12-27 23:44 /tmp  
drwxr-xr-x - hdfs supergroup 0 2017-10-23 09:17 /user  
drwxr-xr-x - hdfs supergroup 0 2017-10-23 09:17 /var
```

```
[cloudera@quickstart ~]$
```

```
hive> create table AgentLoggingReport
```

```
>     (  
>         SL_No float,  
>         Agent varchar(255),  
>         Date date,  
>         Login_Time timestamp,  
>         Logout_Time timestamp,  
>         Duration timestamp  
>     )  
>     row format serde 'org.apache.hadoop.hive.serde2.OpenCSVSerde'  
>     tblproperties ("skip.header.line.count" = "1");
```

OK

Time taken: 0.862 seconds

```
hive> load data inpath '/AgentLoggingReport.csv' into table AgentLoggingReport;
```

Loading data to table hive_class_b1.agentloggingreport

Table hive_class_b1.agentloggingreport stats: [numFiles=1, totalSize=53911]

OK

Time taken: 1.714 seconds

```
hive>
```

Time taken: 1.707 seconds, Fetched: 10 row(s)

```
hive> set hive.cli.print.header = true;
```

```
hive> SELECT * FROM AgentLoggingReport LIMIT 10;
```

OK

```
agentloggingreport.sl_no agentloggingreport.agent agentloggingreport.date agentloggingreport.login_time agentloggingreport.logout_time agentloggingreport.duration  
1 Shivananda Sonwane 30-Jul-22 15:35:29 17:39:39 2:04:10  
2 Khushboo Priya 30-Jul-22 15:06:59 15:07:16 0:00:17  
3 Nandani Gupta 30-Jul-22 15:04:24 17:31:07 2:26:42  
4 Hrisikesh Neogi 30-Jul-22 14:34:29 15:19:35 0:45:06  
5 Mukesh 30-Jul-22 14:03:15 15:11:52 1:08:36  
6 Sowmiya Sivakumar 30-Jul-22 14:03:11 15:05:37 1:02:26  
7 Manjunatha A 30-Jul-22 14:00:12 15:08:29 1:08:16  
8 Harikrishnan Shaji 30-Jul-22 13:53:05 16:06:49 2:13:43  
9 Suraj S Bilgi 30-Jul-22 13:50:01 15:11:42 1:21:41  
10 Shivan K 30-Jul-22 13:28:18 13:59:00 0:30:42
```

Time taken: 0.264 seconds, Fetched: 10 row(s)

```
hive>
```

```

hive> create table AgentPerformance
> (
> SL_No float,
> Date date,
> Agent_Name varchar(255),
> Total_Chats float,
> Average_Response_Time timestamp,
> Average_Resolution_Time timestamp,
> Average_Rating float,
> Total_Feedback float
> )
> row format serde 'org.apache.hadoop.hive.serde2.OpenCSVSerde'
> tblproperties ("skip.header.line.count" = "1");
OK
Time taken: 0.298 seconds
hive> load data inpath '/AgentPerformance.csv' into table AgentPerformance;
Loading data to table hive_class_b1.agentperformance
Table hive_class_b1.agentperformance stats: [numFiles=1, totalSize=112661]
OK
Time taken: 0.878 seconds
hive> 
-----
```

hive> SELECT * FROM AgentPerformance LIMIT 10;

OK

	agentperformance.sl_no	agentperformance.date	agentperformance.agent_name	agentperformance.total_chats	agentperformance.average_response_time	agentperformance.average_resolution_time	agentperformance.average_rating	agentperformance.total_feedback
1	07/30/2022	Prerna Singh	11	0:00:38	0:04:20	4.11	9	
2	07/30/2022	Nandani Gupta	11	0:01:15	0:28:25	3.14	7	
3	07/30/2022	Ameya Jain	14	0:00:30	0:11:36	4.55	11	
4	07/30/2022	Mahesh Sarade	14	0:01:04	0:15:46	4.71	7	
5	07/30/2022	Swati	14	0:01:11	0:16:33	3.67	6	
6	07/30/2022	Mukesh	9	0:00:34	0:15:34	4.62	8	
7	07/30/2022	Saikumarreddy N	15	0:00:31	0:11:09	5	12	
8	07/30/2022	Sanjeev Kumar	12	0:00:55	0:20:43	5	6	
9	07/30/2022	Shubham Sharma	13	0:01:31	0:25:10	4.38	8	
10	07/30/2022	Nishtha Jain	20	0:00:59	0:13:31	4.12	16	

Time taken: 0.31 seconds, Fetched: 10 row(s)

3. List of all agents' names.

```

cloudera@quickstart:~ 
hive> SELECT agent_name FROM AgentLoggingReport
> UNION ALL
> SELECT agent_name FROM AgentPerformance LIMIT 10;
Query ID = cloudera_20230103014646_2e336c8e-8e25-49ad-a62e-e39e5d79d2d2
Total jobs = 2
Launching Job 1 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.reduces=<number>
Starting Job = job_1672736688169_0002, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1672736688169_0002/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1672736688169_0002
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 1
2023-01-03 01:47:04,858 Stage-3 map = 0%, reduce = 0%
2023-01-03 01:47:12,216 Stage-3 map = 100%, reduce = 0%, Cumulative CPU 1.53 sec
2023-01-03 01:47:20,615 Stage-3 map = 100%, reduce = 100%, Cumulative CPU 3.35 sec
MapReduce Total cumulative CPU time: 3 seconds 350 msec
Ended Job = job_1672736688169_0002
Launching Job 2 out of 2
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1672736688169_0003, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1672736688169_0003/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1672736688169_0003
Hadoop job information for Stage-2: number of mappers: 2; number of reducers: 0
2023-01-03 01:47:29,208 Stage-2 map = 0%, reduce = 0%
2023-01-03 01:47:38,745 Stage-2 map = 50%, reduce = 0%, Cumulative CPU 2.65 sec
2023-01-03 01:47:39,792 Stage-2 map = 100%, reduce = 0%, Cumulative CPU 4.91 sec
MapReduce Total cumulative CPU time: 4 seconds 910 msec
Ended Job = job_1672736688169_0003
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1 Reduce: 1 Cumulative CPU: 3.35 sec HDFS Read: 10636 HDFS Write: 391 SUCCESS
Stage-Stage-2: Map: 2 Cumulative CPU: 4.91 sec HDFS Read: 64239 HDFS Write: 13534 SUCCESS
Total MapReduce CPU Time Spent: 8 seconds 260 msec
OK


- _ul.agent_name
- Shivananda Sonwane
- Khushboo Priya
- Nandani Gupta
- Hrisikesh Neogi
- Mukesh
- Sowmiya Sivakumar
- Manjunatha A
- Harikrishnan Shaji

```



```
cloudera@quickstart:~  
_ul.agent_name  
Shivananda Sonwane  
Khushboo Priya  
Nandani Gupta  
Hrisikesh Neogi  
Mukesh  
Sowmiya Sivakumar  
Manjunatha A  
Harikrishnan Shaji  
Suraj S Bilgi  
Shivan K  
Anurag Tiwari  
Ishawant Kumar  
Shivan K  
Shubham Sharma  
Shivan K  
Prerna Singh  
Shivan K  
Shivan K  
Hrisikesh Neogi  
Khushboo Priya  
Nishtha Jain  
Prateek_iot  
Hrisikesh Neogi  
Shivan K  
Nishtha Jain  
Khushboo Priya  
Suraj S Bilgi  
Mithun S  
Harikrishnan Shaji  
Manjunatha A  
Hrisikesh Neogi  
Sowmiya Sivakumar  
Madhulika G  
Boktiar Ahmed Bappy  
Jawala Prakash  
Dibyanshu  
Dibyanshu  
Dibyanshu  
Hrisikesh Neogi  
Dibyanshu  
Boktiar Ahmed Bappy  
Jawala Prakash  
Boktiar Ahmed Bappy  
Jawala Prakash
```

4. Find out agent average rating.

```
cloudera@quickstart:~  
hive> SELECT agent_name, AVG(average_rating) AS Average_Rating FROM Agentperformance GROUP BY agent_name ORDER BY Average_Rating DESC LIMIT 5;  
Query ID = cloudera_20230103015353_7201860f-ab7e-431f-9224-e417547bd8d3  
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.reduces=<number>  
Starting Job = job_1672736688169_0004, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1672736688169_0004/  
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1672736688169_0004  
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1  
2023-01-03 01:53:59,507 Stage-1 map = 0%, reduce = 0%  
2023-01-03 01:54:06,865 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.28 sec  
2023-01-03 01:54:15,272 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 4.11 sec  
MapReduce Total cumulative CPU time: 4 seconds 110 msec  
Ended Job = job_1672736688169_0004  
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.reduces=<number>  
Starting Job = job_1672736688169_0005, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1672736688169_0005/  
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1672736688169_0005  
Hadoop job information for Stage-2: number of mappers: 1; number of reducers: 1  
2023-01-03 01:54:24,380 Stage-2 map = 0%, reduce = 0%  
2023-01-03 01:54:31,827 Stage-2 map = 100%, reduce = 0%, Cumulative CPU 1.96 sec  
2023-01-03 01:54:39,184 Stage-2 map = 100%, reduce = 100%, Cumulative CPU 3.86 sec  
MapReduce Total cumulative CPU time: 3 seconds 860 msec  
Ended Job = job_1672736688169_0005  
MapReduce Jobs Launched:  
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.11 sec HDFS Read: 120841 HDFS Write: 2733 SUCCESS  
Stage-Stage-2: Map: 1 Reduce: 1 Cumulative CPU: 3.86 sec HDFS Read: 7834 HDFS Write: 174 SUCCESS  
Total MapReduce CPU Time Spent: 7 seconds 970 msec  
OK  
agent_name      average_rating  
Shivananda Sonwane      4.2326666666666667  
Khushboo Priya     3.703666666666666  
Manjunatha A       3.5946666666666665  
Boktiar Ahmed Bappy 3.5679999999999996
```

5.Total working days for each agents

```
cloudera@quickstart~  
hive> SELECT agent_name,COUNT(DISTINCT(date)) AS Total_working_days FROM Agentloggingreport GROUP BY agent_name;  
Query ID = cloudera_20230103021919_51dfc441-36e6-44d0-9903-7f5ba57db702  
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_1672736688169_0016, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1672736688169_0016/  
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1672736688169_0016  
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1  
2023-01-03 02:19:29,211 Stage-1 map = 0%, reduce = 0%  
2023-01-03 02:19:35,943 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.65 sec  
2023-01-03 02:19:45,377 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 4.67 sec  
MapReduce Total cumulative CPU time: 4 seconds 670 msec  
Ended Job = job_1672736688169_0016  
MapReduce Jobs Launched:  
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.67 sec HDFS Read: 62695 HDFS Write: 742 SUCCESS  
Total MapReduce CPU Time Spent: 4 seconds 670 msec  
OK  
agent_name      total_working_days  
Aditya Shinde    1  
Aditya_iot       8  
Amersh          2  
Ameya Jain      7  
Ankitjha         2  
Anurag Tiwari   10  
Aravind          7  
Ayushi Mishra   9  
Bharath          8  
Boktiar Ahmed Bappy  9  
Chaitra K Hiremath 7  
Deepranjan Gupta 10  
Dibyanshu        9  
Harikrishnan Shaji 9  
Hrisikesh Neogi  9  
Hyder Abbas     2  
Ineuron Intelligence 1  
Ishawant Kumar   11  
Jawala Prakash   9  
Jaydeep Dixit    7  
Khushboo Priya   8  
Madhulika G      8  
Mahesh Sarade   8  
Maitry           5  
Manjunatha A    7  
Mithun S         8  
Mukesh           2  
Muskan Garg     6  
Nandani Gupta   9  
Nishtha Jain    8  
Nitin M          1  
Prabir Kumar Satapathy 7  
Prateek_iot     11  
Prerna Singh     9  
Rishav Dash     7  
Saikumarreddy N 7  
Sanjeev Kumar   9  
Saurabh Shukla  4  
Shiva Srivastava 8  
Shivan K         8  
Shivananda Sonwane 10  
Shubham Sharma   11  
Sowmiya Sivakumar 8  
Sudhanshu Kumar  6  
Suraj S Bilgi   2  
Swati            4  
Tarun            1  
Wasim            9  
Zeeshan          9  
Time taken: 25.853 seconds, Fetched: 49 row(s)
```

```
cloudera@quickstart~  
Aravind          7  
Ayushi Mishra   9  
Bharath          8  
Boktiar Ahmed Bappy  9  
Chaitra K Hiremath 7  
Deepranjan Gupta 10  
Dibyanshu        9  
Harikrishnan Shaji 9  
Hrisikesh Neogi  9  
Hyder Abbas     2  
Ineuron Intelligence 1  
Ishawant Kumar   11  
Jawala Prakash   9  
Jaydeep Dixit    7  
Khushboo Priya   8  
Madhulika G      8  
Mahesh Sarade   8  
Maitry           5  
Manjunatha A    7  
Mithun S         8  
Mukesh           2  
Muskan Garg     6  
Nandani Gupta   9  
Nishtha Jain    8  
Nitin M          1  
Prabir Kumar Satapathy 7  
Prateek_iot     11  
Prerna Singh     9  
Rishav Dash     7  
Saikumarreddy N 7  
Sanjeev Kumar   9  
Saurabh Shukla  4  
Shiva Srivastava 8  
Shivan K         8  
Shivananda Sonwane 10  
Shubham Sharma   11  
Sowmiya Sivakumar 8  
Sudhanshu Kumar  6  
Suraj S Bilgi   2  
Swati            4  
Tarun            1  
Wasim            9  
Zeeshan          9  
Time taken: 25.853 seconds, Fetched: 49 row(s)
```

```
hive>
```

6. Total query that each agent have taken

```
cloudera@quickstart:~  
agent_name      total_query  
Hrisikesh Neogi 578.0  
Nandani Gupta   560.0  
Zeeshan        542.0  
Maitry         542.0  
Swati          524.0  
Ayushi Mishra  514.0  
Jaydeep Dixit  512.0  
Shubham Sharma 510.0  
Sanjeev Kumar  507.0  
Mithun S       503.0  
Deepranjan Gupta 493.0  
Madhulika G    469.0  
Boktiar Ahmed Bappy 452.0  
Khushboo Priya 446.0  
Shivananda Sonwane 441.0  
Jawala Prakash 439.0  
Wasim          433.0  
Manjunatha A   413.0  
Rishav Dash    409.0  
Prerna Singh   401.0  
Harikrishnan Shaji 381.0  
Nishtha Jain   373.0  
Bharath        369.0  
Aravind        366.0  
Mahesh Sarade  364.0  
Saikumarreddy N 364.0  
Shivan K       357.0  
Ishawant Kumar 338.0  
Ameya Jain    322.0  
Prabir Kumar Satapathy 299.0  
Aditya Shinde  277.0  
Aditya_iot     231.0  
Sowmiya Sivakumar 206.0  
Prateek_iot    190.0  
Jayant Kumar   127.0  
Anirudh        81.0  
Chaitra K Hiremath 64.0  
Muskan Garg    56.0  
Shiva Srivastava 53.0  
Vivek          44.0  
Sandipan Saha  30.0  
Suraj S Bilgi  28.0  
Tarun          22.0  
Mukesh          19.0
```

```
cloudera@quickstart:~  
hive> SELECT agent_name,SUM(Total_chats) AS Total_query FROM Agentperformance GROUP BY agent_name ORDER BY Total_query DESC;  
Query ID = cloudera_20230103023636_ea2b796b-1902-4307-b4a0-4f52b9c20b78  
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.reduces=<number>  
Starting Job = job_1672736688169_0021, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1672736688169_0021/  
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1672736688169_0021  
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1  
2023-01-03 02:36:24,848 Stage-1 map = 0%, reduce = 0%  
2023-01-03 02:36:32,231 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.39 sec  
2023-01-03 02:36:40,612 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 4.22 sec  
MapReduce Total cumulative CPU time: 4 seconds 220 msec  
Ended Job = job_1672736688169_0021  
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.reduces=<number>  
Starting Job = job_1672736688169_0022, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1672736688169_0022/  
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1672736688169_0022  
Hadoop job information for Stage-2: number of mappers: 1; number of reducers: 1  
2023-01-03 02:36:50,270 Stage-2 map = 0%, reduce = 0%  
2023-01-03 02:36:57,647 Stage-2 map = 100%, reduce = 0%, Cumulative CPU 2.26 sec  
2023-01-03 02:37:06,018 Stage-2 map = 100%, reduce = 100%, Cumulative CPU 4.17 sec  
MapReduce Total cumulative CPU time: 4 seconds 170 msec  
Ended Job = job_1672736688169_0022  
MapReduce Jobs Launched:  
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.22 sec HDFS Read: 120569 HDFS Write: 2733 SUCCESS  
Stage-Stage-2: Map: 1 Reduce: 1 Cumulative CPU: 4.17 sec HDFS Read: 7701 HDFS Write: 1228 SUCCESS  
Total MapReduce CPU Time Spent: 8 seconds 390 msec  
OK  
agent_name      total_query  
Hrisikesh Neogi 578.0  
Nandani Gupta   560.0  
Zeeshan        542.0  
Maitry         542.0
```

```
cloudera@quickstart:~  
Type here to search  30°C 4:07 PM  
01/03/2023
```

```
cloudera@quickstart:~  
Ishawant Kumar 338.0  
Ameya Jain 322.0  
Prabir Kumar Satapathy 299.0  
Aditya Shinde 277.0  
Aditya_iot 231.0  
Sowmiya Sivakumar 206.0  
Prateek_iot 190.0  
Jayant Kumar 127.0  
Anirudh 81.0  
Chaitra K Hiremath 64.0  
Muskan Garg 56.0  
Shiva Srivastava 53.0  
Vivek 44.0  
Sandipan Saha 30.0  
Suraj S Bilgi 28.0  
Tarun 22.0  
Mukesh 19.0  
Ashad Nasim 18.0  
Saurabh Shukla 16.0  
Mahak 7.0  
Shivan_S 7.0  
Ankitjha 5.0  
Mukesh Rao 5.0  
Maneesh 4.0  
Anurag Tiwari 4.0  
Sudhanshu Kumar 2.0  
Samprit 1.0  
Hitesh Choudhary 1.0  
Dibyanshu 1.0  
Sanjeevan 0.0  
Saif Khan 0.0  
Rohan 0.0  
Nitin M 0.0  
Spuri 0.0  
Ineuron Intelligence 0.0  
Hyder Abbas 0.0  
Uday Mishra 0.0  
Vasanth P 0.0  
Ashish 0.0  
Ankit Sharma 0.0  
Amersh 0.0  
Aditya 0.0  
Abhishek 0.0  
Time taken: 49.63 seconds, Fetched: 70 row(s)  
hive>
```

7. Total Feedback that each agent have received

```
cloudera@quickstart:~  
hive> SELECT agent_name, SUM(Total_Feedback) AS Total_Feedback_Received FROM Agentperformance GROUP BY agent_name ORDER BY Total_Feedback_Received DESC;  
Query ID = cloudera_20230103024242_3a5dbca5-b891-4c02-9fb6-9489faldc643  
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.reduces=<number>  
Starting Job = job_1672736688169_0024, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1672736688169_0024/  
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1672736688169_0024  
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1  
2023-01-03 02:42:42,924 Stage-1 map = 0%, reduce = 0%  
2023-01-03 02:42:52,044 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.23 sec  
2023-01-03 02:42:59,407 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 5.09 sec  
MapReduce Total cumulative CPU time: 5 seconds 90 msec  
Ended Job = job_1672736688169_0024  
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.reduces=<number>  
Starting Job = job_1672736688169_0025, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1672736688169_0025/  
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1672736688169_0025  
Hadoop job information for Stage-2: number of mappers: 1; number of reducers: 1  
2023-01-03 02:43:08,456 Stage-2 map = 0%, reduce = 0%  
2023-01-03 02:43:16,512 Stage-2 map = 100%, reduce = 0%, Cumulative CPU 2.58 sec  
2023-01-03 02:43:24,899 Stage-2 map = 100%, reduce = 100%, Cumulative CPU 4.6 sec  
MapReduce Total cumulative CPU time: 4 seconds 600 msec  
Ended Job = job_1672736688169_0025  
MapReduce Jobs Launched:  
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 5.09 sec HDFS Read: 120573 HDFS Write: 2733 SUCCESS  
Stage-Stage-2: Map: 1 Reduce: 1 Cumulative CPU: 4.6 sec HDFS Read: 7731 HDFS Write: 1224 SUCCESS  
Total MapReduce CPU Time Spent: 9 seconds 690 msec  
OK  
agent name      total feedback received  
Hrisikesh Neogi 367.0  
Mithun S        364.0  
Maitry          347.0  
Zeeshan          335.0
```

```
cloudera@quickstart:~  
Zeeshan      335.0  
Ayushi Mishra   329.0  
Deepranjan Gupta  312.0  
Boktiar Ahmed Bappy  311.0  
Sanjeev Kumar   311.0  
Nandani Gupta   308.0  
Jaydeep Dixit    305.0  
Swati          302.0  
Shubham Sharma   300.0  
Saikumarreddy N 290.0  
Khushboo Priya   289.0  
Wasim          284.0  
Madhulika G     281.0  
Rishav Dash     264.0  
Shivananda Sonwane 263.0  
Nishtha Jain    257.0  
Manjunatha A     254.0  
Jawala Prakash   250.0  
Bharath          247.0  
Shivan K         243.0  
Prerna Singh     235.0  
Aravind          233.0  
Harikrishnan Shaji 231.0  
Ameya Jain      228.0  
Prabir Kumar Satapathy 222.0  
Mahesh Sarade    216.0  
Ishawant Kumar   202.0  
Aditya Shinde    153.0  
Sowmiya Sivakumar 141.0  
Aditya_iot       131.0  
Prateek_iot      107.0  
Jayant Kumar    70.0  
Shiva Srivastava 46.0  
Anirudh          39.0  
Chaitra K Hiremath 37.0  
Muskan Garg      37.0  
Vivek            20.0  
Sandipan Saha    18.0  
Mukesh           17.0  
Suraj S Bilgi    15.0  
Ashad Nasim      9.0  
Saurabh Shukla   8.0  
Tarun            6.0  
Mahak            5.0  
Mukesh Rao       5.0  
cloudera@quickstart:~
```

```
Type here to search 30°C 4:14 PM 01/03/2023  
cloudera@quickstart:~  
Prabir Kumar Satapathy 222.0  
Mahesh Sarade 216.0  
Ishawant Kumar 202.0  
Aditya Shinde 153.0  
Sowmiya Sivakumar 141.0  
Aditya_iot 131.0  
Prateek_iot 107.0  
Jayant Kumar 70.0  
Shiva Srivastava 46.0  
Anirudh 39.0  
Chaitra K Hiremath 37.0  
Muskan Garg 37.0  
Vivek 20.0  
Sandipan Saha 18.0  
Mukesh 17.0  
Suraj S Bilgi 15.0  
Ashad Nasim 9.0  
Saurabh Shukla 8.0  
Tarun 6.0  
Mahak 5.0  
Mukesh Rao 5.0  
Shivan_S 4.0  
Anurag Tiwari 3.0  
Ankitjha 3.0  
Maneesh 3.0  
Sudhanshu Kumar 2.0  
Samprit 0.0  
Saif Khan 0.0  
Rohan 0.0  
Sanjeevan 0.0  
Nitin M 0.0  
Spuri 0.0  
Ineuron Intelligence 0.0  
Hyder Abbas 0.0  
Hitesh Choudhary 0.0  
Dibyanshu 0.0  
Uday Mishra 0.0  
Vasanth P 0.0  
Ashish 0.0  
Ankit Sharma 0.0  
Amersh 0.0  
Aditya 0.0  
Abhishek 0.0  
Time taken: 50.735 seconds, Fetched: 70 row(s)  
hive> Type here to search 30°C 4:14 PM 01/03/2023
```

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

```
cloudera@quickstart~  
hive> SELECT agent_name,average_rating FROM Agentperformance WHERE average_rating BETWEEN 3.5 AND 4;  
OK  
agent name      average_rating  
Swati          3.67  
Manjunatha A   3.6  
Boktiar Ahmed Bappy  4  
Prateek_iot    3.75  
Nandani Gupta  3.79  
Jaydeep Dixit  3.95  
Mahesh Sarade  3.94  
Zeeshan        3.79  
Hrisikesh Neogi 3.77  
Muskan Garg    4  
Khushboo Priya 3.79  
Wasim          3.95  
Jawala Prakash 3.89  
Shiva Srivastava 4  
Nishtha Jain   3.67  
Maitry         4  
Shiva Srivastava 4  
Nandani Gupta  3.61  
Prerna Singh   3.8  
Shivan K       4  
Boktiar Ahmed Bappy 3.8  
Muskan Garg    3.62  
Chaitra K Hiremath 4  
Harikrishnan Shaji 3.73  
Wasim          3.92  
Deepranjan Gupta 3.69  
Mahesh Sarade  3.81  
Nishtha Jain   3.92  
Shivan K       3.9  
Madhulika G    3.62  
Rishav Dash    3.8  
Prabir Kumar Satapathy 4  
Boktiar Ahmed Bappy 4  
Maitry         3.73  
Nishtha Jain   3.69  
Khushboo Priya 3.95  
Boktiar Ahmed Bappy 3.77  
Mahesh Sarade  3.5  
Rishav Dash    3.67  
Mithun S       3.75  
Jawala Prakash 3.68  
Aditya_iot     3.86
```

```
cloudera@quickstart~  
Sandipan Saha  3.62  
Madhulika G   3.82  
Ayushi Mishra 3.86  
Shivan K      4  
Hrisikesh Neogi 3.8  
Nandani Gupta 3.89  
Harikrishnan Shaji 3.7  
Nishtha Jain   3.67  
Manjunatha A  4  
Anirudh       3.77  
Swati          3.62  
Maitry         3.75  
Madhulika G   4  
Jaydeep Dixit 4  
Zeeshan        3.75  
Nishtha Jain   3.73  
Wasim          3.64  
Manjunatha A  4  
Ishawant Kumar 3.67  
Jayant Kumar   3.76  
Maneesh        4  
Harikrishnan Shaji 4  
Anurag Tiwari 3.5  
Mithun S       3.86  
Prabir Kumar Satapathy 4  
Saurabh Shukla 3.67  
Aditya_iot     3.83  
Maitry         3.92  
Shivananda Sonwane 3.71  
Ameya Jain    4  
Swati          4  
Ashad Nasim   4  
Prabir Kumar Satapathy 4  
Boktiar Ahmed Bappy 3.78  
Nishtha Jain   4  
Ayushi Mishra 3.8  
Shivan K      4  
Shubham Sharma 4  
Jawala Prakash 3.8  
Sanjeev Kumar  4  
Aditya Shinde  3.54  
Deepranjan Gupta 3.71  
Sanjeev Kumar  4  
Time taken: 0.209 seconds, Fetched: 114 row(s)  
hive>
```

cloudera@quickstart~

9. Agent name who have rating less than 3.5

```
cloudera@quickstart~  
hive> SELECT agent_name,average_rating FROM Agentperformance WHERE average_rating < 3.5;  
OK  
agent_name      average_rating  
Nandani Gupta    3.14  
Hitesh Choudhary 0  
Sanjeevan        0  
Anirudh          0  
Shiva Srivastava 0  
Dibyanshu        0  
Ashish           0  
Uday Mishra      0  
Aditya Shinde    0  
Jayant Kumar     0  
Aditya_iot       0  
Prabir Kumar Satapathy 0  
Saurabh Shukla   0  
Wasim            0  
Samprit          0  
Maitry           0  
Abhishek         0  
Rishav Dash      0  
Aravind          0  
Tarun             0  
Muskan Garg      0  
Ankit Sharma     0  
Bharath          0  
Ashad Nasim      0  
Vasanth P         0  
Aditya           0  
Amersh           0  
Sandipan Saha    0  
Ineuron Intelligence 0  
Mahak            0  
Mukesh Rao       0  
Zeeshan          0  
Rishav Dash      0  
Anurag Tiwari    0  
Ankitjha         0  
Saif Khan         0  
Jaydeep Dixit    0  
Sudhanshu Kumar  0  
Maneesh          0  
Hyder Abbas      0  
Ashish           0  
Rohan             0
```

Type here to search 30°C 4:19 PM
cloudera@quickstart~ 01/03/2023 ENG

```
Nitin M 0  
Shivan_S      0  
Spuri         0  
Ayushi Mishra 0  
Chaitra K Hiremath 0  
Prateek_iot   0  
Jawala Prakash 3  
Sowmiya Sivakumar 3.43  
Hitesh Choudhary 0  
Sanjeevan      0  
Anirudh        0  
Dibyanshu      0  
Ashish          0  
Uday Mishra    0  
Aditya Shinde   0  
Nishtha Jain    0  
Jayant Kumar    0  
Swati           0  
Saurabh Shukla 0  
Samprit         0  
Ankit Sharma    0  
Abhishek        0  
Rishav Dash     0  
Aravind          0  
Khushboo Priya 0  
Maitry          0  
Tarun           0  
Ashad Nasim     0  
Vasanth P       0  
Aditya          0  
Amersh          0  
Shubham Sharma  0  
Ineuron Intelligence 0  
Mahak           0  
Sandipan Saha   0  
Mukesh Rao      0  
Manjunatha A    0  
Ameya Jain      0  
Ankitjha        0  
Anurag Tiwari   0  
Saif Khan        0  
Ashish          0  
Harikrishnan Shaji 0  
Sudhanshu Kumar 0  
Maneesh          0
```

Type here to search 30°C 4:20 PM
cloudera@quickstart~ 01/03/2023 ENG

10. Agent name who have rating more than 4.5

```
cloudera@quickstart:~  
hive> SELECT agent_name,average_rating FROM Agentperformance WHERE average_rating > 4.5;  
OK  
agent_name      average_rating  
Ameya Jain      4.55  
Mahesh Sarade   4.71  
Mukesh          4.62  
Saikumarreddy N 5  
Sanjeev Kumar   5  
Harikrishnan Shaji 4.57  
Sowmiya Sivakumar 4.75  
Boktiar Ahmed Bappy 4.75  
Shivananda Sonwane 5  
Ishawant Kumar  4.67  
Deepranjan Gupta 4.8  
Shivananda Sonwane 4.67  
Muskan Garg     5  
Aditya_iot      4.6  
Mukesh          4.67  
Chaitra K Hiremath 5  
Suraj S Bilgi   5  
Mithun S        4.8  
Wasim           4.71  
Bharath          4.85  
Ishawant Kumar  5  
Jawala Prakash  4.83  
Ayushi Mishra   5  
Ameya Jain      4.64  
Bharath          4.79  
Ishawant Kumar  4.75  
Shivananda Sonwane 5  
Ishawant Kumar  4.88  
Prateek_iot     4.83  
Sanjeev Kumar   4.82  
Nandani Gupta   4.62  
Saikumarreddy N 4.59  
Ameya Jain      5  
Khushboo Priya  4.71  
Wasim           4.53  
Manjunatha A    4.54  
Bharath          4.88  
Maitry          5  
Ayushi Mishra   5  
Deepranjan Gupta 5  
Muskan Garg     4.6  
Prateek_iot     5
```

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

```
cloudera@quickstart:~  
hive> SELECT agent_name,SUM(total_feedback) AS Total_feedbacks FROM Agentperformance WHERE average_rating > 4.5 GROUP BY agent_name ORDER BY Total_feedbacks DESC;  
Query ID = cloudera_20230103032828_e846fe2e-d1e1-4042-911c-04b988d73a9d  
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.reducer.bytes.per.reducer=<number>  
In order to limit the maximum number of reducers:  
  set hive.exec.reducer.max=<number>  
In order to set a constant number of reducers:  
  set mapreduce.job.reduces=<number>  
Starting Job = job_1672736688169_0027, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1672736688169_0027/  
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1672736688169_0027  
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1  
2023-01-03 03:29:02,941 Stage-1 map = 0%, reduce = 0%  
2023-01-03 03:29:12,762 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.12 sec  
2023-01-03 03:29:21,150 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 5.04 sec  
MapReduce Total cumulative CPU time: 5 seconds 40 msec  
Ended Job = job_1672736688169_0027  
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.reducer.bytes.per.reducer=<number>  
In order to limit the maximum number of reducers:  
  set hive.exec.reducer.max=<number>  
In order to set a constant number of reducers:  
  set mapreduce.job.reduces=<number>  
Starting Job = job_1672736688169_0028, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1672736688169_0028/  
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1672736688169_0028  
Hadoop job information for Stage-2: number of mappers: 1; number of reducers: 1  
2023-01-03 03:29:30,125 Stage-2 map = 0%, reduce = 0%  
2023-01-03 03:29:36,407 Stage-2 map = 100%, reduce = 0%, Cumulative CPU 1.41 sec  
2023-01-03 03:29:44,780 Stage-2 map = 100%, reduce = 100%, Cumulative CPU 3.38 sec  
MapReduce Total cumulative CPU time: 3 seconds 380 msec  
Ended Job = job_1672736688169_0028  
MapReduce Jobs Launched:  
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 5.04 sec HDFS Read: 121402 HDFS Write: 1896 SUCCESS  
Stage-Stage-2: Map: 1 Reduce: 1 Cumulative CPU: 3.38 sec HDFS Read: 6878 HDFS Write: 863 SUCCESS  
Total MapReduce CPU Time Spent: 8 seconds 420 msec  
OK  
agent_name      total_feedbacks  
Bharath          231.0  
Saikumarreddy N 184.0  
Hrisikesh Neogi 183.0
```

cloudera@quickstart:~

```
agent_name      total_feedbacks
Bharath        231.0
Saikumarreddy N 184.0
Hrisikesh Neogi 183.0
Jaydeep Dixit   179.0
Aravind         178.0
Sanjeev Kumar   164.0
Wasim           156.0
Shubham Sharma  155.0
Shivananda Sonwane 154.0
Ameya Jain     150.0
Khushboo Priya 134.0
Manjunatha A   132.0
Shivan K        130.0
Swati           103.0
Mithun S        93.0
Nandani Gupta   91.0
Deepranjan Gupta 83.0
Prerna Singh    82.0
Maitry          81.0
Harikrishnan Shaji 80.0
Ishawant Kumar  79.0
Ayushi Mishra   75.0
Aditya Shinde   73.0
Madhulika G     67.0
Prateek_iot     54.0
Prabir Kumar Satapathy 53.0
Boktiar Ahmed Bappy 52.0
Zeeshan          48.0
Mahesh Sarade   46.0
Aditya_iot       43.0
Jawala Prakash  33.0
Rishav Dash     27.0
Nishtha Jain    18.0
Mukesh           17.0
Sowmiya Sivakumar 16.0
Muskan Garg     14.0
Jayant Kumar    7.0
Chaitra K Hiremath 5.0
Suraj S Bilgi   4.0
Saurabh Shukla  3.0
Vivek            3.0
Shiva Srivastava 2.0
Anirudh          2.0
Sudhanshu Kumar 2.0
```

5:00 PM 01/03/2023

Hive> 

```
All rows          2.0
Sudhanshu Kumar 2.0
Sandipan Saha   1.0
Ankitjha        1.0
Mukesh Rao      1.0
Time taken: 50.566 seconds, Fetched: 47 row(s)
```

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

```
cloudera@quickstart~  
hive> SELECT agent_name,SUM(Total_chats) AS Total_chats FROM Agentperformance WHERE total_feedback NOT IN (0) GROUP BY agent_name;  
Query ID = cloudera_20230103035454_98c4866c-820f-404f-ba4d-4846fdded2b  
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_1672736688169_0032, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1672736688169_0032/  
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1672736688169_0032  
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1  
2023-01-03 03:54:39,990 Stage-1 map = 0%, reduce = 0%  
2023-01-03 03:54:58,200 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 7.73 sec  
2023-01-03 03:55:12,734 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 12.32 sec  
MapReduce Total cumulative CPU time: 12 seconds 320 msec  
Ended Job = job_1672736688169_0032  
MapReduce Jobs Launched:  
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 12.32 sec HDFS Read: 122385 HDFS Write: 977 SUCCESS  
Total MapReduce CPU Time Spent: 12 seconds 320 msec  
OK  
agent_name      total_chats  
Aditya Shinde    277.0  
Aditya_iot       231.0  
Ameya Jain      322.0  
Anirudh         80.0  
Ankitjha        4.0  
Anurag Tiwari   4.0  
Aravind          366.0  
Ashad Nasim     18.0  
Ayushi Mishra   514.0  
Bharath          369.0  
Boktiar Ahmed Bappy 450.0  
Chaitra K Hiremath 64.0  
Deepranjan Gupta 492.0  
Harikrishnan Shaji 381.0  
Hrisikesh Neogi  578.0  
Ishawant Kumar   338.0  
Jawala Prakash   439.0  
Jayant Kumar     127.0  
Jaydeep Dixit    512.0  
Khushboo Priya   444.0  
Madhulika G      469.0
```

```
cloudera@quickstart~  
Boktiar Ahmed Bappy 450.0  
Chaitra K Hiremath 64.0  
Deepranjan Gupta 492.0  
Harikrishnan Shaji 381.0  
Hrisikesh Neogi 578.0  
Ishawant Kumar 338.0  
Jawala Prakash 439.0  
Jayant Kumar 127.0  
Jaydeep Dixit 512.0  
Khushboo Priya 444.0  
Madhulika G 469.0  
Mahak 7.0  
Mahesh Sarade 364.0  
Maitry 542.0  
Maneesh 3.0  
Manjunatha A 411.0  
Mithun S 503.0  
Mukesh 19.0  
Mukesh Rao 5.0  
Muskan Garg 56.0  
Nandani Gupta 560.0  
Nishtha Jain 373.0  
Prabir Kumar Satapathy 299.0  
Prateek_iot 189.0  
Prerna Singh 401.0  
Rishav Dash 407.0  
Saikumarreddy N 364.0  
Sandipan Saha 30.0  
Sanjeev Kumar 507.0  
Saurabh Shukla 16.0  
Shiva Srivastava 53.0  
Shivan K 355.0  
Shivan_S 7.0  
Shivananda Sonwane 441.0  
Shubham Sharma 510.0  
Sowmiya Sivakumar 206.0  
Sudhanshu Kumar 2.0  
Suraj S Bilgi 28.0  
Swati 524.0  
Tarun 22.0  
Vivek 44.0  
Wasim 433.0  
Zeeshan 542.0  
Time taken: 49.225 seconds, Fetched: 53 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.

cloudera@quickstart:~

```
hive> create table innerjoin_agent
>      (
>        sl_no string,
>        date string,
>        agent string,
>        login_time string,
>        logout_time string,
>        duration string,
>        sl_no2 string,
>        date2 string,
>        agent_name string,
>        total_chats string,
>        average_response_time string,
>        average_resolution_time string,
>        average_rating string,
>        total_feedback string)
>      row format delimited
>      fields terminated by ',';

OK
Time taken: 0.663 seconds
hive> insert overwrite table tbl_innerjoinagent select * from agentloggingreport AL inner join agentperformance AP on AL.agent_name = AP.agent_name;
FAILED: SemanticException [Error 10001]: Line 1:23 Table not found 'tbl_innerjoinagent'
hive> insert overwrite table innerjoin_agent select * from agentloggingreport AL inner join agentperformance AP on AL.agent_name = AP.agent_name;
Query ID = cloudera_20230103044747_45d00215-6f2e-4885-8bcf-c466a6c4e699
Total jobs = 1
Execution log at: /tmp/cloudera/cloudera_20230103044747_45d00215-6f2e-4885-8bcf-c466a6c4e699.log
2023-01-03 04:48:06      Starting to launch local task to process map join;      maximum memory = 932184064
2023-01-03 04:48:11      Dump the side-table for tag: 0 with group count: 49 into file: file:/tmp/cloudera/c61dlbba-5233-4bb8-9980-16adec6adc7f/hive_2023-01-03_04-47-51_660_3567203862881472802-1/-local-10002/HashTable-Stage-4/MapJoin-mapfile00--.hashtable
2023-01-03 04:48:12      Uploaded 1 File to: file:/tmp/cloudera/c61dlbba-5233-4bb8-9980-16adec6adc7f/hive_2023-01-03_04-47-51_660_3567203862881472802-1/-local-10002/HashTable-Stage-4/MapJoin-mapfile00--.hashtable (46210 bytes)
2023-01-03 04:48:12      End of local task; Time Taken: 5.529 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_1672736688169_0033, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1672736688169_0033/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1672736688169_0033
Hadoop job information for Stage-4: number of mappers: 1; number of reducers: 0
2023-01-03 04:48:31,929 Stage-4 map = 0%, reduce = 0%
2023-01-03 04:48:50,980 Stage-4 map = 100%, reduce = 0%, Cumulative CPU 9.4 sec
MapReduce Total cumulative CPU time: 9 seconds 400 msec
Ended Job = job_1672736688169_0033
Loading data to table hive_class_b1.innerjoin_agent
Table hive_class_b1.innerjoin_agent stats: [numFiles=1, numRows=21180, totalSize=2313915, rawDataSize=2292735]

6:19 PM 01/03/2023

hive> set hive.cli.print.header = true;
hive> SELECT * FROM innerjoin_agent LIMIT 2;
OK
innerjoin_agent.sl_no    innerjoin_agent.date    innerjoin_agent.agent    innerjoin_agent.login_time    innerjoin_agent.logout_time    innerjoin_agent.duration
innerjoin_agent.sl_no2   innerjoin_agent.date2   innerjoin_agent.agent_name   innerjoin_agent.total_chats   innerjoin_agent.average_response_time
innerjoin_agent.average_resolution_time innerjoin_agent.average_rating innerjoin_agent.total_feedback
16      Prerna Singh    30-Jul-22      12:32:28      14:10:08      1:37:40 1      07/30/2022      Prerna Singh    11      0:00:38 0:04:20 4.11    9
75      Prerna Singh    29-Jul-22      17:47:06      21:03:44      3:16:37 1      07/30/2022      Prerna Singh    11      0:00:38 0:04:20 4.11    9
Time taken: 0.165 seconds, Fetched: 2 row(s)
hive>

hive> create table leftjoin_agent
>      (
>        sl_no string,
>        date string,
>        agent string,
>        login_time string,
>        logout_time string,
>        duration string)
>      row format delimited
>      fields terminated by ',';

OK
Time taken: 0.234 seconds
```

```

hive> insert overwrite table leftjoin_agent select AL.sl_no,AL.date,AL.agent_name,AL.login_time,AL.logout_time,AL.duration from agentloggingreport as AL left join agentperformance as AP on AL.agent_name = AP.agent_name;
Query ID = cloudera_20230103045454_c26182ae-98e1-40f4-8ebb-e3295588ee58
Total jobs = 1
Execution log at: /tmp/cloudera/cloudera_20230103045454_c26182ae-98e1-40f4-8ebb-e3295588ee58.log
2023-01-03 04:54:19 Starting to launch local task to process map join; maximum memory = 932184064
2023-01-03 04:54:22 Dump the side-table for tag: 1 with group count: 70 into file: file:/tmp/cloudera/c61dlbba-5233-4bb8-9980-16adec6adc7f/hive_2023-01-03_04-54-08_030_8950828534475427491-1/-local-10002/HashTable-Stage-4/MapJoin-mapfile11--.hashtable
2023-01-03 04:54:22 Uploaded 1 File to: file:/tmp/cloudera/c61dlbba-5233-4bb8-9980-16adec6adc7f/hive_2023-01-03_04-54-08_030_8950828534475427491-1/-local-10002/HashTable-Stage-4/MapJoin-mapfile11--.hashtable (10727 bytes)
2023-01-03 04:54:22 End of local task; Time Taken: 2.938 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_1672736688169_0034, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1672736688169_0034/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1672736688169_0034
Hadoop job information for Stage-4: number of mappers: 1; number of reducers: 0
2023-01-03 04:54:38,704 Stage-4 map = 0%, reduce = 0%
2023-01-03 04:54:55,225 Stage-4 map = 100%, reduce = 0%, Cumulative CPU 7.12 sec
MapReduce Total cumulative CPU time: 7 seconds 120 msec
Ended Job = job_1672736688169_0034
Loading data to table hive_class_b1.leftjoin_agent
Table hive_class_b1.leftjoin_agent stats: [numFiles=1, numRows=21486, totalSize=1170090, rawDataSize=1148604]
MapReduce Jobs Launched:
Stage-Stage-4: Map: 1 Cumulative CPU: 7.12 sec HDFS Read: 62023 HDFS Write: 1170182 SUCCESS
Total MapReduce CPU Time Spent: 7 seconds 120 msec
OK
al.sl_no      al.date al.agent_name   al.login_time   al.logout_time   al.duration
Time taken: 50.342 seconds
hive>

```

```

hive> SELECT * FROM leftjoin_agent LIMIT 5;
OK
leftjoin_agent.sl_no    leftjoin_agent.date    leftjoin_agent.agent    leftjoin_agent.login_time    leftjoin_agent.logout_time    leftjoin_agent.duration
n
1      30-Jul-22    Shivananda Sonwane    15:35:29    17:39:39    2:04:10
Time taken: 0.169 seconds, Fetched: 5 row(s)
hive>

```

```

cloudera@quickstart:~$ 
hive> create table rightjoin_agentperformance
>     (
>         sl_no string,
>         date string,
>         agent_name string,
>         total_chats string,
>         average_response_time string,
>         average_resolution_time string,
>         average_rating string,
>         total_feedback string)
>         row format delimited
>         fields terminated by ',';
OK
Time taken: 0.221 seconds

```

```

hive> insert overwrite table rightjoin_agentperformance select AP.sl_no,AP.date,AP.agent_name,AP.total_chats,AP.average_response_time string,AP.average_resolution_time string,AP.average_rating string,AP.total_feedback string from agentloggingreport as AL right join agentperformance as AP on AL.agent_name = AP.agent_name;
Query ID = cloudera_20230103050505_c45858be-9fc8-4b86-bb4f-cf4f15a6b0bc
Total jobs = 1
Execution log at: /tmp/cloudera/cloudera_20230103050505_c45858be-9fc8-4b86-bb4f-cf4f15a6b0bc.log
2023-01-03 05:06:08 Starting to launch local task to process map join; maximum memory = 932184064
2023-01-03 05:06:11 Dump the side-table for tag: 0 with group count: 49 into file: file:/tmp/cloudera/c61d1bba-5233-4bb8-9980-16adec6adc7f/hive_2023-01-03_05-05-57_265_5348707461005635033-1-local-10002/HashTable-Stage-4/MapJoin-mapfile20--.hashtable
2023-01-03 05:06:11 Uploaded 1 File to: file:/tmp/cloudera/c61d1bba-5233-4bb8-9980-16adec6adc7f/hive_2023-01-03_05-05-57_265_5348707461005635033-1-local-10002/HashTable-Stage-4/MapJoin-mapfile20--.hashtable (5560 bytes)
2023-01-03 05:06:11 End of local task; Time Taken: 2.924 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_1672736688169_0035, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1672736688169_0035/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1672736688169_0035
Hadoop job information for Stage-4: number of mappers: 1; number of reducers: 0
2023-01-03 05:06:30,566 Stage-4 map = 0%, reduce = 0%
2023-01-03 05:06:47,201 Stage-4 map = 100%, reduce = 0%, Cumulative CPU 7.82 sec
MapReduce Total cumulative CPU time: 7 seconds 820 msec
Ended Job = job_1672736688169_0035
Loading data to table hive_class_b1.rightjoin_agentperformance
Table hive_class_b1.rightjoin_agentperformance stats: [numFiles=1, numRows=22260, totalSize=1210855, rawDataSize=1188595]
MapReduce Jobs Launched:
Stage-Stage-4: Map: 1 Cumulative CPU: 7.82 sec HDFS Read: 121204 HDFS Write: 1210959 SUCCESS
Total MapReduce CPU Time Spent: 7 seconds 820 msec
OK
ap.sl_no      ap.date ap.agent_name  ap.total_chats  string  string  string  string
Time taken: 51.977 seconds
hive> 
Time taken: 51.977 seconds
hive> SELECT * FROM rightjoin_agentperformance LIMIT 5;
OK
rightjoin_agentperformance.sl_no      rightjoin_agentperformance.date rightjoin_agentperformance.agent_name  rightjoin_agentperformance.total_chats  rightjoin_agentperformance.average_response_time      rightjoin_agentperformance.average_resolution_time      rightjoin_agentperformance.average_rating      rightjoin_agentperformance.total_feedback
1      07/30/2022      Prerna Singh    11      0:00:38 0:04:20 4.11      9
1      07/30/2022      Prerna Singh    11      0:00:38 0:04:20 4.11      9
1      07/30/2022      Prerna Singh    11      0:00:38 0:04:20 4.11      9
1      07/30/2022      Prerna Singh    11      0:00:38 0:04:20 4.11      9
1      07/30/2022      Prerna Singh    11      0:00:38 0:04:20 4.11      9
Time taken: 0.149 seconds, Fetched: 5 row(s)
hive> 
```

```

drwxrwxrwx  - cloudera supergroup          0 2022-09-06 04:08 /user/hive/warehouse/hive_class_b1.db/department_data_hdfs
drwxrwxrwx  - cloudera supergroup          0 2023-01-03 04:48 /user/hive/warehouse/hive_class_b1.db/innerjoin_agent
drwxrwxrwx  - cloudera supergroup          0 2023-01-03 04:54 /user/hive/warehouse/hive_class_b1.db/leftjoin_agentloggingreport
drwxrwxrwx  - cloudera supergroup          0 2022-09-20 00:11 /user/hive/warehouse/hive_class_b1.db/locations
drwxrwxrwx  - cloudera supergroup          0 2022-09-29 03:53 /user/hive/warehouse/hive_class_b1.db/parking_violations_issued_fiscalyear
drwxrwxrwx  - cloudera supergroup          0 2023-01-03 05:06 /user/hive/warehouse/hive_class_b1.db/rightjoin_agentperformance
drwxrwxrwx  - cloudera supergroup          0 2022-09-12 23:38 /user/hive/warehouse/hive_class_b1.db/sales_data
[cloudera@quickstart ~]$ hdfs dfs -cp /user/hive/warehouse/hive_class_b1.db/innerjoin_agent 'file:///tmp/hive_class_local_tmp'
[cloudera@quickstart ~]$ cd /tmp/hive_class_local_tmp
[cloudera@quickstart hive_class_local_tmp]$ ls
AgentLoggingReport.csv  AgentPerformance.csv  CustomersCSV.csv  dept_csv.txt  innerjoin_agent  sales_order_data_csv.txt
[cloudera@quickstart hive_class_local_tmp]$ ls innerjoin_agent
000000 0_
[cloudera@quickstart ~]$ hdfs dfs -cp /user/hive/warehouse/hive_class_b1.db/leftjoin_agentloggingreport 'file:///tmp/hive_class_local_tmp'
[cloudera@quickstart ~]$ hdfs dfs -cp /user/hive/warehouse/hive_class_b1.db/rightjoin_agentperformance 'file:///tmp/hive_class_local_tmp'
[cloudera@quickstart ~]$ cd /tmp/hive_class_local_tmp
[cloudera@quickstart hive_class_local_tmp]$ ls
AgentLoggingReport.csv  CustomersCSV.csv  innerjoin_agent          rightjoin_agentperformance
AgentPerformance.csv   dept_csv.txt     leftjoin_agentloggingreport  sales_order_data_csv.txt
[cloudera@quickstart hive_class_local_tmp]$ 
```

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

```
cloudera@quickstart~  
hive> set hive.exec.dynamic.partition.mode=nonstrict;  
hive> create table agentloggingreport_dynamic_part  
>     (br/>>     sl_no string,  
>     date string,  
>     login_time string,  
>     logout_time string,  
>     duration string)  
>     partitioned by (agent string);  
OK  
Time taken: 0.189 seconds  
hive> insert overwrite table agentloggingreport_dynamic_part partition(agent) select sl_no,date,login_time,logout_time,duration,agent from agentloggingreport;  
FAILED: SemanticException [Error 10004]: Line 1:121 Invalid table alias or column reference 'agent': (possible column names are: sl_no, agent_name, date, login_time, logout_time, duration)  
hive> insert overwrite table agentloggingreport_dynamic_part partition(agent) select sl_no,date,login_time,logout_time,duration,agent_name from agentloggingreport;  
Query ID = cloudera_20230103090303_4860dfca-d94a-4910-a07f-e77500d2786f  
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_1672762349728_0004, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1672762349728_0004/  
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1672762349728_0004  
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 0  
2023-01-03 09:03:26,753 Stage-1 map = 0%, reduce = 0%  
2023-01-03 09:03:35,134 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.04 sec  
MapReduce Total cumulative CPU time: 3 seconds 40 msec  
Ended Job = job_1672762349728_0004  
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_class_b1.db/agentloggingreport_dynamic_part/.hive-staging_hive_2023-01-03_09-03-18_414  
_8578061315494461567-1-ext-10000  
Loading data to table hive_class_b1.agentloggingreport_dynamic_part partition (agent=null)  
    Time taken for load dynamic partitions : 10061  
    Loading partition {agent=Tarun}  
    Loading partition {agent=Prateek_iot}  
    Loading partition {agent=Nandani_Gupta}  
    Loading partition {agent=Sudhanshu_Kumar}  
    Loading partition {agent=Rishav_Dash}  
    Loading partition {agent=Muskan_Garg}  
    Loading partition {agent=Prabir_Kumar_Satapathy}  
    Loading partition {agent=Harikrishnan_Shaji}  
    Loading partition {agent=Swati}  
    Loading partition {agent=Aditya_Shinde}  
    Loading partition {agent=Maitry}  
Type here to search 24°C 10:34 PM  
cloudera@quickstart~ 01/03/2023 ENG
```

```
Stage-Stage-1: Map: 1 Cumulative CPU: 3.04 sec HDFS Read: 59233 HDFS Write: 43621 SUCCESS  
Total MapReduce CPU Time Spent: 3 seconds 40 msec  
OK  
Time taken: 31.828 seconds  
hive> set hive.enforce.bucketing=true;  
hive> create table buck_agentloggingreport_dynamic_part  
>     (br/>>     sl_no float,  
>     date date,  
>     agent string,  
>     login_time timestamp,  
>     logout_time timestamp,  
>     duration timestamp)  
>     clustered by (agent)  
>     sorted by (agent)  
>     into 2 buckets;  
OK  
Time taken: 0.088 seconds  
hive> insert overwrite table buck_agentloggingreport_dynamic_part select * from agentloggingreport_dynamic_part;  
Query ID = cloudera_20230103090707_6ebald01-e2a5-4966-9c29-b6f7fb3967de  
Total jobs = 1  
Launching Job 1 out of 1  
Number of reduce tasks determined at compile time: 2  
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_1672762349728_0005, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1672762349728_0005/  
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1672762349728_0005  
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 2  
2023-01-03 09:07:13,257 Stage-1 map = 0%, reduce = 0%  
2023-01-03 09:07:19,657 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.14 sec  
2023-01-03 09:07:28,126 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 8.3 sec  
MapReduce Total cumulative CPU time: 8 seconds 300 msec  
Ended Job = job_1672762349728_0005  
Loading data to table hive_class_b1.buck_agentloggingreport_dynamic_part  
Table hive_class_b1.buck_agentloggingreport_dynamic_part stats: [numFiles=2, numRows=1000, totalSize=26657, rawDataSize=25657]  
MapReduce Jobs Launched:  
Stage-Stage-1: Map: 1 Reduce: 2 Cumulative CPU: 8.3 sec HDFS Read: 77932 HDFS Write: 26875 SUCCESS  
Total MapReduce CPU Time Spent: 8 seconds 300 msec  
OK  
Time taken: 24.778 seconds  
hive> Type here to search 24°C 10:37 PM  
cloudera@quickstart~ 01/03/2023 ENG
```

 cloudera@quickstart:~

```
[cloudera@quickstart ~]$ hdfs dfs -ls /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part
Found 49 items
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Aditya Shinde
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Aditya_iot
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Amersh
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Ameya Jain
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Ankitjha
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Anurag Tiwari
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Aravind
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Ayushi Mishra
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Bharath
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Boktiar Ahmed Bappy
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Chaitra K Hiremath
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Deepranjan Gupta
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Dibyanshu
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Harikrishnan Shaji
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Hrisikesh Neogi
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Hyder Abbas
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Ineuron Intelligence
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Ishawant Kumar
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Jawala Prakash
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Jaydeep Dixit
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Khushboo Priya
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Madhulika G
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Mahesh Sarade
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Maitry
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Manjunatha A
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Mithun S
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Mukesh
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Muskan Garg
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Nandani Gupta
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Nishtha Jain
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Nitin M
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Prabir Kumar Satapat
hy
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Prateek_iot
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Prerna Singh
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Rishav Dash
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Saikumarreddy N
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Sanjeev Kumar
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Saurabh Shukla
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Shiva Srivastava
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Shivan K
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Shivananda Sonwane
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Shubham Sharma
```

```
[cloudera@quickstart ~]$ hdfs dfs -ls /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Wasim
Found 1 items
-rwxrwxrwx  1 cloudera supergroup      793 2023-01-03 09:03 /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Wasim/000000_0
[cloudera@quickstart ~]$ hdfs dfs -cat /user/hive/warehouse/hive_class_b1.db/agentloggingreport_dynamic_part/agent=Wasim/000000_0
13129-Jul-229:00:0716:19:577:19:50
19528-Jul-2216:04:5216:07:120:02:20
22628-Jul-229:23:3215:35:406:12:07
35427-Jul-229:28:0014:52:055:24:05
45726-Jul-2211:20:5411:22:440:01:50
45826-Jul-2211:20:5415:06:433:45:49
50025-Jul-2217:49:2317:51:440:02:21
50825-Jul-2216:23:3116:24:140:00:43
51025-Jul-2215:38:0415:40:040:02:00
53125-Jul-2212:08:5915:05:552:56:56
54325-Jul-229:14:4211:59:542:45:12
62323-Jul-2223:02:4423:03:390:00:55
74222-Jul-2215:38:4915:40:490:02:00
80822-Jul-229:23:2114:55:515:32:30
86721-Jul-2216:09:3016:15:000:05:30
88721-Jul-229:17:3015:20:316:03:00
91320-Jul-2223:20:1023:20:380:00:27
94320-Jul-2219:42:5220:48:131:05:20
96220-Jul-2216:13:1416:54:340:41:20
99120-Jul-229:17:2515:23:566:06:30
[cloudera@quickstart ~]$
```

```
[cloudera@quickstart ~]$ hdfs dfs -ls /user/hive/warehouse/hive class b1.db/buck agentloggingreport dynamic part
```

[Header at top]

```
[cloudera@quickstart ~]$ hdfs dfs -ls /user/hive/warehouse/hive_class_b1.db/buck_agentloggingreport_dynamic_part/000000_0  
-rwxrwxrwx 1 cloudera supergroup 13901 2023-01-03 09:07 /user/hive/warehouse/hive_class_b1.db/buck_agentloggingreport_dynamic_part/000000_0  
-rwxrwxrwx 1 cloudera supergroup 12756 2023-01-03 09:07 /user/hive/warehouse/hive_class_b1.db/buck_agentloggingreport_dynamic_part/000001_0
```

```
cloudera@quickstart:~
```

```
hive> set hive.exec.dynamic.partition.mode=nonstrict;
hive> create table agentperformance_dynamic_part
  >   (
  >     sl_no string,
  >     date string,
  >     total_chats string,
  >     average_response_time string,
  >     average_resolution_time string,
  >     average_rating string,
  >     total_feedback string)
  >   partitioned by (agent_name string);
OK
Time taken: 0.392 seconds
hive> insert overwrite table agentperformance_dynamic_part partition(agent_name) select sl_no,date,total_chats,average_response_time,average_resolution_time,average_rating string,total_feedback string,agent_name from agentperformance;
Query ID = cloudera_20230103091515_b686af4b-4b5c-42ce-8772-96053b195941
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_1672762349728_0006, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1672762349728_0006/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1672762349728_0006
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 0
2023-01-03 09:15:53,763 Stage-1 map = 0%, reduce = 0%
2023-01-03 09:16:04,373 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 4.18 sec
MapReduce Total cumulative CPU time: 4 seconds 180 msec
Ended Job = job_1672762349728_0006
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_class_b1.db/agentperformance_dynamic_part/.hive-staging_hive_2023-01-03_09-15-44_655_
6081700429380515584-1/-ext-10000
Loading data to table hive_class_b1.agentperformance_dynamic_part partition (agent_name=null)
  Time taken for load dynamic partitions : 11556
  Loading partition {agent_name=Shivan_S }
  Loading partition {agent_name=Nandani Gupta}
  Loading partition {agent_name=Sanjeev Kumar}
  Loading partition {agent_name=Jayant Kumar}
  Loading partition {agent_name=Prateek_iot }
  Loading partition {agent_name=Suraj S Bilgi}
  Loading partition {agent_name=Jawala Prakash}
  Loading partition {agent_name=Ishawant Kumar}
  Loading partition {agent_name=Jaydeep Dixit}
  Loading partition {agent_name=Aditya }
  Loading partition {agent_name=Ameya Jain}
  Loading partition {agent_name=Tarun }
```

```
cloudera@quickstart:~
```

```
Type here to search 24°C 10:46 PM
01/03/2023
cloudera@quickstart:~
Time taken: 37.114 seconds
hive> set hive.enforce.bucketing=true;
hive> create table buck_agentperformance_dynamic_part
  >   (
  >     sl_no string,
  >     date string,
  >     agent_name string,
  >     total_chats string,
  >     average_response_time string,
  >     average_resolution_time string,
  >     average_rating string,
  >     total_feedback string)
  >   clustered by(agent_name)
  >   sorted by (agent_name)
  >   into 3 buckets;
OK
Time taken: 0.086 seconds
hive> insert overwrite table buck_agentperformance_dynamic_part select * from agentperformance_dynamic_part;
Query ID = cloudera_20230103091818_3bba8401-b127-437c-86cd-3739d4f2078a
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 3
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_1672762349728_0007, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1672762349728_0007/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1672762349728_0007
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 3
2023-01-03 09:18:21,514 Stage-1 map = 0%, reduce = 0%
2023-01-03 09:18:31,097 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.28 sec
2023-01-03 09:18:45,098 Stage-1 map = 100%, reduce = 33%, Cumulative CPU 6.38 sec
2023-01-03 09:18:46,146 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 13.07 sec
MapReduce Total cumulative CPU time: 13 seconds 70 msec
Ended Job = job_1672762349728_0007
Loading data to table hive_class_b1.buck_agentperformance_dynamic_part
Table hive_class_b1.buck_agentperformance_dynamic_part stats: [numFiles=3, numRows=2160, totalSize=110390, rawDataSize=108230]
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 3 Cumulative CPU: 13.07 sec HDFS Read: 137791 HDFS Write: 110715 SUCCESS
Total MapReduce CPU Time Spent: 13 seconds 70 msec
OK
Time taken: 33.801 seconds
hive>
```

```
cloudera@quickstart:~
```

cloudera@quickstart:~

```
[cloudera@quickstart ~]$ hdfs dfs -ls /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part
Found 70 items
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Abhishek
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Aditya
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Aditya Shinde
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Aditya_iot
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Amersh
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Ameya Jain
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Anirudh
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Ankit Sharma
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Ankitjha
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Anurag Tiwari
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Aravind
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Ashad Nasim
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Ashish
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Ayushi Mishra
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Bharath
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Boktiar Ahmed Ba
ppy
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Chaitra K Hirema
th
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Deepranjan Gupta
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Dibyanshu
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Harikrishnan Sha
ji
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Hitesh Choudhary
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Hrisikesh Neogi
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Hyder Abbas
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Ineuron Intelligence
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Ishawant Kumar
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Jawala Prakash
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Jayant Kumar
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Jaydeep Dixit
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Khushboo Priya
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Madhulika G
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Mahak
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Mahesh Sarade
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Maitry
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Maneesh
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Manjunatha A
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Mithun S
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Mukesh
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Mukesh Rao
drwxrwxrwx - cloudera supergroup          0 2023-01-03 09:16 /user/hive/warehouse/hive_class_b1.db/agentperformance_dynamic_part/agent_name=Muskan Garg
```

Type here to search                                                     <img alt="Apache Beam icon" data-bbox="993 3901