1. Create a schema based on the given dataset

```
hive> show databases;
OK
agentdata
default
orders
Time taken: 2.283 seconds, Fetched: 3 row(s)
hive> use agentdata
    > ;
OK
Time taken: 0.114 seconds
hive> create table loginReport(
    > SL_NO int,
    > AGENT string,
    > DATE_l string,
    > LOGIN_TIME string,
    > LOGOUT_TIME string,
    > DURATION string
    > )
    > row format delimited
    > fields terminated by ','
    > tblproperties ("skip.header.line.count"="1");
OK
Time taken: 0.425 seconds
hive> create table performance(
    > SL_NO int,
    > DATE_l string,
    > AGENT_NAME string,
    > TOTAL_CHATS int,
    > AVERAGE_RESPONSE_TIME string,
    > AVERAGE_RESOLUTION_TIME string,
    > AVERAGE_RATING float,
    > TOTAL_FEEDBACK int)
    > row format delimited
    > fields terminated by ','
    > tblproperties ("skip.header.line.count"="1");
OK
Time taken: 0.13 seconds
hive>
```

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

```
OK
Time taken: 0.13 seconds
hive> load data inpath "/tmp/agentdata/AgentLogingReport.csv" into table loginReport;
Loading data to table agentdata.loginreport
Table agentdata.loginreport stats: [numFiles=1, totalSize=55351]
OK
Time taken: 0.895 seconds
hive> load data inpath "/tmp/agentdata/AgentPerformance.csv" into table performance;
Loading data to table agentdata.performance
Table agentdata.performance stats: [numFiles=1, totalSize=109853]
OK
Time taken: 0.357 seconds
hive> set hive.cli.header=true;
hive> select * from loginReport limit 2;
OK
1       Shivananda Sonwane      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
Time taken: 0.8 seconds, Fetched: 2 row(s)
hive> select * from performance limit 2;
OK
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
Time taken: 0.151 seconds, Fetched: 2 row(s)
hive>
```

3. List of all agents' names.

```
hive> select agent from loginReport limit 20;
OK
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
Time taken: 0.097 seconds, Fetched: 20 row(s)
hive>
```

4. Find out agent average rating.

```
hive> select agent_name,average_rating from performance limit 20;
OK
Prerna Singh    4.11
Nandani Gupta   3.14
Ameya Jain      4.55
Mahesh Sarade   4.71
Swati   3.67
Mukesh  4.62
Saikumarreddy N 5.0
Sanjeev Kumar   5.0
Shubham Sharma  4.38
Nishtha Jain    4.12
Manjunatha A    3.6
Khushboo Priya  4.43
Suraj S Bilgi   4.36
Harikrishnan Shaji      4.57
Hrisikesh Neogi 4.3
Shivan K        4.17
Sowmiya Sivakumar       4.75
Madhulika G     4.25
Mithun S        4.05
Hitesh Choudhary        0.0
Time taken: 0.101 seconds, Fetched: 20 row(s)
hive>
```

5. Total working days for each agents

```
hive> select agent,count(*)  as working_days from loginReport
    > group by agent
    > limit 20;
Query ID = cloudera_20221030071919_197f63d4-f084-451e-a51e-4574276d55a7
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_1663302417861_0039, Tracking URL = http://quickstart.cloudera:8088/proxy/application_16633024178
_0039/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1663302417861_0039
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-30 07:20:18,057 Stage-1 map = 0%,  reduce = 0%
2022-10-30 07:20:36,182 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 4.47 sec
2022-10-30 07:20:53,744 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 9.74 sec
MapReduce Total cumulative CPU time: 9 seconds 740 msec
Ended Job = job_1663302417861_0039
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 9.74 sec   HDFS Read: 63756 HDFS Write: 328 SUCCESS
Total MapReduce CPU Time Spent: 9 seconds 740 msec
OK
Aditya Shinde   1
Aditya_iot      9
Amersh  4
Ameya Jain      10
Ankitjha        4
Anurag Tiwari   37
Aravind 10
Ayushi Mishra   18
Bharath 9
Boktiar Ahmed Bappy     17
Chaitra K Hiremath      13
Deepranjan Gupta        58
Dibyanshu       208
Harikrishnan Shaji      23
Hrisikesh Neogi 37
Hyder Abbas     2
Ineuron Intelligence    1
Ishawant Kumar  49
Jawala Prakash  16
Jaydeep Dixit   11
Time taken: 58.837 seconds, Fetched: 20 row(s)
hive>
```

6. Total query that each agent have taken

```
hive> select agent_name,sum(total_chats)  as query from performance
    > group by agent_name
    > limit 20;
Query ID = cloudera_20221030072525_6ffb4a32-df3f-4145-a3e6-f397cdff4238
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_1663302417861_0041, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1663
_0041/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1663302417861_0041
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-30 07:26:16,911 Stage-1 map = 0%,  reduce = 0%
2022-10-30 07:26:34,521 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 4.67 sec
2022-10-30 07:26:55,700 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 11.46 sec
MapReduce Total cumulative CPU time: 11 seconds 460 msec
Ended Job = job_1663302417861_0041
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 11.46 sec   HDFS Read: 119013 HDFS Write: 308 SUCCESS
Total MapReduce CPU Time Spent: 11 seconds 460 msec
OK
Abhishek        0
Aditya  0
Aditya Shinde   277
Aditya_iot      231
Amersh  0
Ameya Jain      322
Anirudh         81
Ankit Sharma    0
Ankitjha        5
Anurag Tiwari   4
Aravind         366
Ashad Nasim     18
Ashish  0
Ayushi Mishra   514
Bharath         369
Boktiar Ahmed Bappy     452
Chaitra K Hiremath      64
Deepranjan Gupta        493
Dibyanshu       1
Harikrishnan Shaji      381
Time taken: 61.159 seconds, Fetched: 20 row(s)
hive>
```

7. Total Feedback that each agent have receive.

```
hive> select agent_name,sum(total_feedback) as Total_Feedback
    > from performance
    > group by agent_name
    > limit 20;
Query ID = cloudera_20221030073131_6372fd31-52e0-4021-8674-123f934f295a
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_1663302417861_0042, Tracking URL = http://quickstart.cloudera:8088/proxy/applic
_0042/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1663302417861_0042
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-30 07:31:37,410 Stage-1 map = 0%,  reduce = 0%
2022-10-30 07:31:52,740 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 4.58 sec
2022-10-30 07:32:12,255 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 9.81 sec
MapReduce Total cumulative CPU time: 9 seconds 810 msec
Ended Job = job_1663302417861_0042
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 9.81 sec   HDFS Read: 119014 HDFS Write: 307 SU
Total MapReduce CPU Time Spent: 9 seconds 810 msec
OK
Abhishek        0
Aditya  0
Aditya Shinde   153
Aditya_iot      131
Amersh  0
Ameya Jain      228
Anirudh         39
Ankit Sharma    0
Ankitjha        3
Anurag Tiwari   3
Aravind         233
Ashad Nasim     9
Ashish  0
Ayushi Mishra   329
Bharath         247
Boktiar Ahmed Bappy     311
Chaitra K Hiremath      37
Deepranjan Gupta        312
Dibyanshu       0
Harikrishnan Shaji      231
Time taken: 60.952 seconds, Fetched: 20 row(s)
```

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

```
hive> select agent_name,average_rating from performance
    > where average_rating between 3.5 AND 4
    > limit 20;
OK
Swati   3.67
Manjunatha A    3.6
Boktiar Ahmed Bappy     4.0
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     4.0
Khushboo Priya  3.79
Wasim   3.95
Jawala Prakash  3.89
Shiva Srivastava        4.0
Nishtha Jain    3.67
Maitry  4.0
Shiva Srivastava        4.0
Nandani Gupta   3.61
Prerna Singh    3.8
Shivan K        4.0
Time taken: 0.104 seconds, Fetched: 20 row(s)
hive>
```

9. Agent name who have rating less than 3.5

```
hive> select agent_name,average_rating from performance
    > where average_rating < 3.5
    > limit 20;
OK
Nandani Gupta   3.14
Hitesh Choudhary        0.0
Sanjeevan       0.0
Anirudh         0.0
Shiva Srivastava        0.0
Dibyanshu       0.0
Ashish  0.0
Uday Mishra     0.0
Aditya Shinde   0.0
Jayant Kumar    0.0
Aditya_iot      0.0
Prabir Kumar Satapathy  0.0
Saurabh Shukla  0.0
Wasim   0.0
Samprit         0.0
Maitry  0.0
Abhishek        0.0
Rishav Dash     0.0
Aravind         0.0
Tarun   0.0
Time taken: 0.262 seconds, Fetched: 20 row(s)
hive>
```

10. Agent name who have rating more than 4.5

```
hive> select agent_name,average_rating from performance
    > where average_rating > 4.5
    > limit 20;
OK
Ameya Jain      4.55
Mahesh Sarade   4.71
Mukesh  4.62
Saikumarreddy N 5.0
Sanjeev Kumar   5.0
Harikrishnan Shaji      4.57
Sowmiya Sivakumar       4.75
Boktiar Ahmed Bappy     4.75
Shivananda Sonwane      5.0
Ishawant Kumar  4.67
Deepranjan Gupta        4.8
Shivananda Sonwane      4.67
Muskan Garg     5.0
Aditya_iot      4.6
Mukesh  4.67
Chaitra K Hiremath      5.0
Suraj S Bilgi   5.0
Mithun S        4.8
Wasim   4.71
Bharath         4.85
Time taken: 0.201 seconds, Fetched: 20 row(s)
hive>
```

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

```
hive> select date_l,agent_name,average_response_time as weekly_response from performance
    > group by date_l,agent_name,average_response_time
    > limit 10;
Query ID = cloudera_20221030083232_45936666-b51b-4f00-807c-fb4302d7cbfa
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_1663302417861_0048, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1663
_0048/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1663302417861_0048
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-30 08:33:05,392 Stage-1 map = 0%,  reduce = 0%
2022-10-30 08:33:22,461 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 5.51 sec
2022-10-30 08:33:41,063 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 11.16 sec
MapReduce Total cumulative CPU time: 11 seconds 160 msec
Ended Job = job_1663302417861_0048
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 11.16 sec   HDFS Read: 119099 HDFS Write: 279 SUCCESS
Total MapReduce CPU Time Spent: 11 seconds 160 msec
OK
7/1/2022        Abhishek        0:00:00
7/1/2022        Aditya  0:00:00
7/1/2022        Aditya Shinde   0:00:00
7/1/2022        Aditya_iot      0:00:00
7/1/2022        Amersh  0:00:00
7/1/2022        Ameya Jain      0:00:00
7/1/2022        Anirudh         0:00:00
7/1/2022        Ankit Sharma    0:00:00
7/1/2022        Ankitjha        0:00:00
7/1/2022        Anurag Tiwari   0:00:00
Time taken: 57.782 seconds, Fetched: 10 row(s)
hive>
```

7/8/2022	Tarun	0:00:00
7/8/2022	Uday Mishra	0:00:00
7/8/2022	Vasanth P	0:00:00
7/8/2022	Vivek	0:01:00
7/8/2022	Wasim	0:00:00
7/8/2022	Zeeshan	0:01:23
7/9/2022	Abhishek	0:00:00
7/9/2022	Aditya	0:00:00
7/9/2022	Aditya Shinde	0:01:14
7/9/2022	Aditya_iot	0:00:00
7/9/2022	Amersh	0:00:00
7/9/2022	Ameya Jain	0:00:00
7/9/2022	Anirudh	0:05:10
7/9/2022	Ankit Sharma	0:00:00
7/9/2022	Ankitjha	0:00:00
7/9/2022	Anurag Tiwari	0:00:00
7/9/2022	Aravind	0:00:39
7/9/2022	Ashad Nasim	0:00:00
7/9/2022	Ashish	0:00:00
7/9/2022	Ayushi Mishra	0:01:34
7/9/2022	Bharath	0:00:25
7/9/2022	Boktiar Ahmed Bappy	0:00:57
7/9/2022	Chaitra K Hiremath	0:00:00
7/9/2022	Deepranjan Gupta	0:01:07
7/9/2022	Dibyanshu	0:00:00
7/9/2022	Harikrishnan Shaji	0:00:43
7/9/2022	Hitesh Choudhary	0:00:00
7/9/2022	Hrisikesh Neogi	0:00:40
7/9/2022	Hyder Abbas	0:00:00
7/9/2022	Ineuron Intelligence	0:00:00
7/9/2022	Ishawant Kumar	0:00:25
7/9/2022	Jawala Prakash	0:00:00
7/9/2022	Jayant Kumar	0:00:00
7/9/2022	Jaydeep Dixit	0:01:15
7/9/2022	Khushboo Priya	0:00:35
7/9/2022	Madhulika G	0:01:29
7/9/2022	Mahak	0:00:00
7/9/2022	Mahesh Sarade	0:01:41
7/9/2022	Maitry	0:00:00
7/9/2022	Maneesh	0:00:39
7/9/2022	Manjunatha A	0:00:38
7/9/2022	Mithun S	0:00:00
7/9/2022	Mukesh	0:00:00
7/9/2022	Mukesh Rao	0:00:00
7/9/2022	Muskan Garg	0:00:00
7/9/2022	Nandani Gupta	0:01:01
7/9/2022	Nishtha Jain	0:00:57
7/9/2022	Nitin M	0:00:00
7/9/2022	Prabir Kumar Satapathy	0:01:36
7/9/2022	Prateek _iot	0:00:00

12. average weekly response time for each agent

```
hive> select agent_name,sum(total_feedback) as Feedback_count from performance
    > group by agent_name
    > limit 20;
Query ID = cloudera_20221030083838_82138a71-6349-44eb-8a34-e3e58bbbce1c
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_1663302417861_0050, Tracking URL = http://quickstart.cloudera:8088/proxy/applica
_0050/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1663302417861_0050
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-10-30 08:38:29,781 Stage-1 map = 0%,  reduce = 0%
2022-10-30 08:38:47,072 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 6.34 sec
2022-10-30 08:39:06,790 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 12.33 sec
MapReduce Total cumulative CPU time: 12 seconds 330 msec
Ended Job = job_1663302417861_0050
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 12.33 sec   HDFS Read: 119028 HDFS Write: 307 SU
Total MapReduce CPU Time Spent: 12 seconds 330 msec
OK
Abhishek        0
Aditya  0
Aditya Shinde   153
Aditya_iot      131
Amersh  0
Ameya Jain      228
Anirudh         39
Ankit Sharma    0
Ankitjha        3
Anurag Tiwari   3
Aravind         233
Ashad Nasim     9
Ashish  0
Ayushi Mishra   329
Bharath         247
Boktiar Ahmed Bappy     311
Chaitra K Hiremath      37
Deepranjan Gupta        312
Dibyanshu       0
Harikrishnan Shaji      231
Time taken: 59.887 seconds, Fetched: 20 row(s)
hive>
```

cloudera@quickstart:~/Desktop