1. Create a schema based on the given dataset

```
Time taken: 0.012 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 ',';
OK
Time taken: 0.077 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 ',';
OK
Time taken: 0.08 seconds
hive>
```

2.

```
Time taken: 0.069 seconds
hive> load data local inpath 'file:///tmp/AgentLogingR
FAILED: SemanticException Line 1:23 Invalid path ''file
hive> load data local inpath 'file:///tmp/data/AgentLo
Loading data to table db1.loginreport
Table db1.loginreport stats: [numFiles=1, totalSize=55
OK
Time taken: 0.669 seconds
hive> load data local inpath 'file:///tmp/data/AgentPe
Loading data to table db1.performance
Table db1.performance stats: [numFiles=1, totalSize=10
OK
Time taken: 0.241 seconds
hive>
```

3. List of all agents' names.

```
hive> select Agent from LoginReport;
OK
Agent
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
```

4. Find out agent average rating.

```
hive> select Average_Rating from Performance limit 20;
OK
NULL
4.11
3.14
4.55
4.71
3.67
4.62
5.0
5.0
4.38
4.12
3.6
4.43
4.36
4.57
4.3
4.17
4.75
4.25
4.05
Time taken: 0.06 seconds, Fetched: 20 row(s)
hive>
```

5. Total working days for each agents

```
FAILED: ParseException line 1:05 missing EOF at 'by' hea
hive> select count(*) as days from Performance group by(
FAILED: SemanticException [Error 10004]: Line 1:50 Inval
tal_chats, average_response_time, average_resolution_tim
hive> select count(*) as days from Performance group by(
FAILED: SemanticException [Error 10004]: Line 1:50 Inval
tal_chats, average_response_time, average_resolution_tim
hive> select count(*) as days from LoginReport group by(
Query ID = cloudera_20220928110909_c6cf124b-aa15-40b4-b3
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks not specified. Estimated from inp
In order to change the average load for a reducer (in by
  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_1664363535866_0006, Tracking URL = ht
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job
Hadoop job information for Stage-1: number of mappers: 1
2022-09-28 11:09:32,510 Stage-1 map = 0%,   reduce = 0%
2022-09-28 11:09:40,267 Stage-1 map = 100%,   reduce = 0%
2022-09-28 11:09:50,153 Stage-1 map = 100%,   reduce = 10
MapReduce Total cumulative CPU time: 2 seconds 260 msec
Ended Job = job_1664363535866_0006
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 2.26
Total MapReduce CPU Time Spent: 2 seconds 260 msec
OK
1
9
1
4
```

```
18
9
17
13
58
208
23
37
2
1
49
16
11
18
17
36
5
8
14
3
12
11
18
1
26
17
18
12
10
20
40
15
36
15
35
24
11
5
5
1
20
10
Time taken: 27.8 seconds, Fetched: 50 row(s)
hive>
```

6. Total query that each agent have taken

7. Total Feedback that each agent have received

```
ve> select count(Total_Chats) as query from Performance gr
ery ID = cloudera_20220928112424_2ef9c685-6ff4-41fd-a30d-d
tal jobs = 1
```

```
OK
query
30
30
30
30
0
30
30
30
30
30
30
30
60
30
30
30
30
30
30
30
30
30
30
30
30
30
30
```

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

```
Time taken: 30.266 seconds, Fetched: 71 row(s)
hive> select Agent_Name from  Performance where Average_Rating BETWEEN 3.5 and 4;
OK
agent_name
Swati
Manjunatha A
Boktiar Ahmed Bappy
Prateek _iot
Nandani Gupta
Jaydeep Dixit
Mahesh Sarade
Zeeshan
Hrisikesh Neogi
Muskan Garg
Khushboo Priya
Wasim
Jawala Prakash
Shiva Srivastava
Nishtha Jain
Maitry
Shiva Srivastava
Nandani Gupta
Prerna Singh
Shivan K
Boktiar Ahmed Bappy
Muskan Garg
Chaitra K Hiremath
Harikrishnan Shaji
Wasim
Deepranjan Gupta
Mahesh Sarade
Nishtha Jain
Shivan K
Madhulika G
Rishav Dash
Prabir Kumar Satapathy
Boktiar Ahmed Bappy
Maitry
Nishtha Jain
Khushboo Priya
Boktiar Ahmed Bappy
Mahesh Sarade
Rishav Dash
```

9. Agent name who have rating less than 3.5

```
Time taken: 0.179 seconds, Fetched: 114 row(s)
hive> select Agent_Name from  Performance where Average_Rating < 3.5;
OK
agent_name
Nandani Gupta
Hitesh Choudhary
Sanjeevan
Anirudh
Shiva Srivastava
Dibyanshu
Ashish
Uday Mishra
Aditya Shinde
Jayant Kumar
Aditya_iot
Prabir Kumar Satapathy
Saurabh Shukla
Wasim
Samprit
Maitry
Abhishek
Rishav Dash
Aravind
Tarun
Muskan Garg
Ankit Sharma
Bharath
Ashad Nasim
```

10. Agent name who have rating more than 4.5

```
Time taken: 0.100 seconds, Fetched: 1171 row(s)
hive> select Agent_Name from  Performance where Average_Rating > 4.5;
OK
agent_name
Ameya Jain
Mahesh Sarade
Mukesh
Saikumarreddy N
Sanjeev Kumar
Harikrishnan Shaji
Sowmiya Sivakumar
Boktiar Ahmed Bappy
Shivananda Sonwane
Ishawant Kumar
Deepranjan Gupta
Shivananda Sonwane
Muskan Garg
Aditya_iot
Mukesh
Chaitra K Hiremath
Suraj S Bilgi
Mithun S
Wasim
Bharath
Ishawant Kumar
Jawala Prakash
Ayushi Mishra
```

12. average weekly response time for each agent

```
0:01:21
0:01:02
0:01:51
0:00:24
0:00:40
0:00:47
0:00:32
0:01:49
0:00:51
0:00:38
0:00:59
0:00:44
0:00:45
0:01:01
0:00:30
0:00:25
0:00:39                 0:00:00
0:00:29
0:00:53          0:00:00
0:01:05
0:01:28          Time taken: 0.061 seconds, Fetched: 2161 row(s)
0:01:43
0:04:02          hive> select average_response_time from Performance;
0:00:00
```

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

```
ime taken: 0.395 seconds
ive>  select sum(total_feedback) as total  from Performance group by agent_name;
uery ID = cloudera_20220928123434_f890c70f-021b-4872-973e-3e4a9f92afa1
otal jobs = 1
aunching Job 1 out of 1
umber of reduce tasks not specified. Estimated from input data size: 1
n order to change the average load for a reducer (in bytes):
 set hive.exec.reducers.bytes.per.reducer=<number>
```

```
Total MapReduce CPU Time Spent: 2 seconds 410 msec
OK
0
0
153
131
NULL
0
228
39
0
3
3
233
9
0
329
247
311
37
312
0
231
0
367
0
0
202
250
70
305
```

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.

```
1000        Tarun      01-Jul-22          01:52:47          12:01:07              514:08:20         1000
Time taken: 30.403 seconds, Fetched: 1000 row(s)
hive> SELECT * FROM loginreport u INNER JOIN performance l ON u.SL_NO=l.SL_NO;
Query ID = cloudera_20220928125757_251e1b0d-cae0-403c-a50e-bcd4c3acad44
Total jobs = 1
```

```
[cloudera@quickstart ~]$ vim export.hql
[cloudera@quickstart ~]$ hive -f 'export.hql' > localfile.txt
```

```
OK
Time taken: 30.055 seconds, Fetched: 1001 row(s)
[cloudera@quickstart ~]$ cat localfile.txt
NULL    Agent   Date    Login Time      Logout Time     Duration        NULL    NULL
1       Shivananda Sonwane      30-Jul-22       15:35:29        17:39:39        02:04:10
20      4.11    9
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
1
4       Hrisikesh Neogi 30-Jul-22       14:34:29        15:19:35        00:45:06
5       Mukesh  30-Jul-22       14:03:15        15:11:52        01:08:36        5
6       Sowmiya Sivakumar       30-Jul-22       14:03:11        15:05:37        01:02:2
7       Manjunatha A    30-Jul-22       14:00:12        15:08:29        01:08:16
2
8       Harikrishnan Shaji      30-Jul-22       13:53:05        16:06:49        02:13:4
43      5.0     6
9       Suraj S Bilgi   30-Jul-22       13:50:01        15:11:42        01:21:41
10      Shivan K        30-Jul-22       13:28:18        13:59:00        00:30:42
6
11      Anurag Tiwari   30-Jul-22       13:06:12        13:11:57        00:05:44
12      Ishawant Kumar  30-Jul-22       13:05:35        13:12:45        00:07:10
4
13      Shivan K        30-Jul-22       13:01:33        13:27:53        00:26:20
1
14      Shubham Sharma  30-Jul-22       12:48:50        13:03:10        00:14:20
49      4.57    7
15      Shivan K        30-Jul-22       12:34:27        12:40:37        00:06:10
0
16      Prerna Singh    30-Jul-22       12:32:28        14:10:08        01:37:40
2
```

```
1000    Tarun   01-Jul-22       01:32:47        12:01:07        314.08.20       1000
Time taken: 30.32 seconds, Fetched: 1000 row(s)
hive> SELECT * FROM loginreport u LEFT JOIN performance l ON u.SL_NO=l.SL_NO;
Query ID = cloudera_20220928130000_197cd918-bf11-406c-8abc-6586072c2a2f
Total jobs = 1
Execution log at: /tmp/cloudera/cloudera_20220928130000_197cd918-bf11-406c-8abc-6586072
2022-09-28 01:00:14     Starting to launch local task to process map join;      maximum
2022-09-28 01:00:16     Dump the side-table for tag: 1 with group count: 2161 into file
-28_13-00-09_455_6051060040595615669-1/-local-10003/HashTable-Stage-3/MapJoin-mapfile21
2022-09-28 01:00:16     Uploaded 1 File to: file:/tmp/cloudera/4efae792-f986-42b6-b685-
-10003/HashTable-Stage-3/MapJoin-mapfile21--.hashtable (141083 bytes)
2022-09-28 01:00:16     End of local task; Time Taken: 2.05 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_1664363535866_0015, Tracking URL = http://quickstart.cloudera:8088/p
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1664363535866_0015
```

```
Time taken: 29.623 seconds, Fetched: 1001 row(s)
hive> SELECT * FROM loginreport u RIGHT JOIN performance l ON u.SL_NO=l.SL_NO;
Query ID = cloudera_20220928130202_4d27d509-e089-4622-a06c-bd351bdcd8af
Total jobs = 1
```

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

```
Total MapReduce CPU Time Spent: 0 msec
hive> set hive.exec.dynamic.partition.mode=nonstrict;
hive> insert overwrite table LoginReport_partion partition
Query ID = cloudera_20221001015353_31f1b083-365e-40f1-bcdb
Total jobs = 3
Launching Job 1 out of 3
Number of reduce tasks is set to 0 since there's no reduce
Starting Job = job_1664363535866_0020, Tracking URL = http
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1
Hadoop job information for Stage-1: number of mappers: 1;
2022-10-01 01:53:15,732 Stage-1 map = 0%,   reduce = 0%
2022-10-01 01:53:25,906 Stage-1 map = 100%,   reduce = 0%
```

```
cloudera@quickstart:~

        Loading partition {agent=Agent}
        Loading partition {agent=Shivan K}
        Loading partition {agent=Shiva Srivastava}
        Loading partition {agent=Saurabh Shukla}
        Loading partition {agent=Aditya_iot}
        Loading partition {agent=Ayushi Mishra}
        Loading partition {agent=Sudhanshu Kumar}
        Loading partition {agent=Jawala Prakash}
        Loading partition {agent=Shivananda Sonwane}
        Loading partition {agent=Sowmiya Sivakumar}
        Loading partition {agent=Maitry}
        Loading partition {agent=Khushboo Priya}
        Loading partition {agent=Hrisikesh Neogi}
        Loading partition {agent=Ineuron Intelligence}
        Loading partition {agent=Anurag Tiwari}
        Loading partition {agent=Muskan Garg}
        Loading partition {agent=Aravind}
        Loading partition {agent=Jaydeep Dixit}
        Loading partition {agent=Nishtha Jain}
        Loading partition {agent=Deepranjan Gupta}
        Loading partition {agent=Mahesh Sarade}
        Loading partition {agent=Prerna Singh}
        Loading partition {agent=Mithun S}
        Loading partition {agent=Ankitjha}
        Loading partition {agent=Hyder Abbas}
        Loading partition {agent=Shubham Sharma}
        Loading partition {agent=Swati}
        Loading partition {agent=Nandani Gupta}
        Loading partition {agent=Saikumarreddy N}
        Loading partition {agent=Mukesh}
        Loading partition {agent=Bharath}
         Time taken for adding to write entity : 23
Partition db1.loginreport_partion{agent=Aditya Shinde} sta
Partition db1.loginreport_partion{agent=Aditya_iot} stats:
Partition db1.loginreport_partion{agent=Agent} stats: [num
Partition db1.loginreport_partion{agent=Amersh} stats: [nu
Partition db1.loginreport_partion{agent=Ameya Jain} stats:
Partition db1.loginreport_partion{agent=Ankitjha} stats: [
Partition db1.loginreport_partion{agent=Anurag Tiwari} sta
Partition db1.loginreport_partion{agent=Aravind} stats: [n
Partition db1.loginreport_partion{agent=Ayushi Mishra} sta
Partition db1.loginreport_partion{agent=Bharath} stats: [n
Partition db1.loginreport_partion{agent=Boktiar Ahmed Bapp
Partition db1.loginreport_partion{agent=Chaitra K Hiremath
```

```
Time taken: 0.08 seconds, Fetched: 13 row(s)
hive> insert overwrite table  performance_partion  partiti
ge_rating,total_feedback,agent_name from performance;
Query ID = cloudera_20221001021818_319b719d-337d-4936-a43e
Total jobs = 3
Launching Job 1 out of 3
```

```
Stage-5 is filtered out by condition resolver.
Moving data to: hdfs://quickstart.cloudera:8020/user/hive/
1/-ext-10000
Loading data to table db1.performance_partion partition (a
        Time taken for load dynamic partitions : 10796
      Loading partition {agent_name=Swati }
      Loading partition {agent_name=Boktiar Ahmed Bappy}
      Loading partition {agent_name=Suraj S Bilgi}
      Loading partition {agent_name=Shivan_S }
      Loading partition {agent_name=Ashad Nasim}
      Loading partition {agent_name=Harikrishnan Shaji}
      Loading partition {agent_name=Prerna Singh}
      Loading partition {agent_name=Chaitra K Hiremath}
      Loading partition {agent_name=Hrisikesh Neogi}
      Loading partition {agent_name=Deepranjan Gupta}
      Loading partition {agent_name=Madhulika G}
      Loading partition {agent_name=Abhishek }
      Loading partition {agent_name=Sanjeevan }
      Loading partition {agent_name=Rishav Dash}
      Loading partition {agent_name=Aditya_iot }
      Loading partition {agent_name=Agent Name}
      Loading partition {agent_name=Sanjeev Kumar}
      Loading partition {agent_name=Amersh }
      Loading partition {agent_name=Nitin M}
      Loading partition {agent_name=Saurabh Shukla}
      Loading partition {agent_name=Ayushi Mishra}
      Loading partition {agent_name=Manjunatha A}
      Loading partition {agent_name=Aditya Shinde}
      Loading partition {agent_name=Wasim }
      Loading partition {agent_name=Ankit Sharma}
      Loading partition {agent_name=Mahesh Sarade}
      Loading partition {agent_name=Shiva Srivastava}
      Loading partition {agent_name=Zeeshan }
      Loading partition {agent_name=Uday Mishra}
      Loading partition {agent_name=Mukesh Rao }
      Loading partition {agent_name=Sudhanshu Kumar}
      Loading partition {agent_name=Vasanth P}
      Loading partition {agent_name=Maitry }
      Loading partition {agent_name=Hyder Abbas}
      Loading partition {agent_name=Ashish }
      Loading partition {agent_name=Anirudh }
      Loading partition {agent_name=Ishawant Kumar}
      Loading partition {agent_name=Jaydeep Dixit}
      Loading partition {agent_name=Saif Khan}
```