HIVE PROJECT 1

Q1) Create a schema based on the given dataset

Q2) Dump the data inside the hdfs in the given schema location.

Q1,2) Solution

Make directory at hdfs location.

[cloudera@quickstart ~]$ hdfs dfs -mkdir Agent

Put files (AgentLogingReport.csv Agent, AgentPerformance.csv Agent) into Agent directory in hdfs location.

[cloudera@quickstart ~]$ hdfs dfs -put /tmp/Agent_data/AgentLogingReport.csv Agent/

[cloudera@quickstart ~]$ hdfs dfs -put /tmp/Agent_data/AgentPerformance.csv Agent/

Check the files inside hdfs location.

[cloudera@quickstart ~]$ hdfs dfs -ls Agent

Found 2 items

-rw-r--r--   1 cloudera cloudera   55351 2022-11-01 08:58 Agent/AgentLogingReport.csv

-rw-r--r--   1 cloudera cloudera   109853 2022-11-01 08:59 Agent/AgentPerformance.csv

Go to hive and use agent database.

hive> use agent;

OK

Time taken: 1.25 seconds

Create tables:

hive> create table agent_loging

    > (

    > s_no int,

    > agent string,

    > date date,

    > login_time string,

    > logout_time string,

    > duration string

    >)

    > row format delimited

    > fields terminated by ','

    > tblproperties("skip.header.line.count"="1");

OK

Time taken: 0.662 seconds

hive> create table agent_performance

   > (

   > s_no int,

   > date date,

   > agent_name string,

   > total_chats int,

   > average_response_time string,

   > average_resolution_time string,

   > average_rating float,

   > total_feedback int

   >)

   > row format delimited

   > fields terminated by ','

   > tblproperties("skip.header.line.count"="1");

OK

Time taken: 0.218 seconds

Load data inside above created tables.

hive> load data inpath 'Agent/AgentLogingReport.csv' into table agent_loging;

Loading data to table agent.agent_loging

Table agent.agent_loging stats: [numFiles=1, totalSize=56353]

OK

Time taken: 1.581 seconds

hive> load data inpath 'Agent/AgentPerformance.csv' into table agent_performance;

Loading data to table agent.agent_performance

Table agent.agent_performance stats: [numFiles=1, totalSize=116159]

OK

Time taken: 0.837 seconds

Fetch some records from tables.

OK

1	Shivananda Sonwane	2022-07-30	15:35:29	17:39:39	02:04:10
2	Khushboo Priya	2022-07-30	15:06:59	15:07:16	00:00:17
3	Nandani Gupta	2022-07-30	15:04:24	17:31:07	02:26:42

Time taken: 0.971 seconds, Fetched: 3 row(s)

hive> select * from agent_performance limit 3;

OK

1	2022-07-30	Prerna Singh	11	00:00:38	00:04:20	4.11	9
2	2022-07-30	Nandani Gupta	11	00:01:15	00:28:25	3.14	7
3	2022-07-30	Ameya Jain	14	00:00:30	00:11:36	4.55	11

Time taken: 0.118 seconds, Fetched: 3 row(s)

Q3) List of all agents' names.

Solution:

hive> select distinct agent from agent_loging;

Query ID = cloudera_20221101210404_30dc2792-6a5d-46e6-95d7-61106454adf0

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_1667316033891_0002, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667316033891_0002/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667316033891_0002

Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1

2022-11-01 21:05:08,000 Stage-1 map = 0%, reduce = 0%

2022-11-01 21:05:23,698 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.81 sec

2022-11-01 21:05:37,446 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6.11 sec

MapReduce Total cumulative CPU time: 6 seconds 110 msec

Ended Job = job_1667316033891_0002

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 6.11 sec   HDFS Read: 64055 HDFS Write: 638 SUCCESS

Total MapReduce CPU Time Spent: 6 seconds 110 msec

OK

Aditya Shinde

Aditya_iot

Amersh

Ameya Jain

Ankitjha

Anurag Tiwari

Aravind

Ayushi Mishra

Bharath

Boktiar Ahmed Bappy

Chaitra K Hiremath

Deepranjan Gupta

Dibyanshu

Harikrishnan Shaji

Hrisikesh Neogi

Hyder Abbas

Ineuron Intelligence

Ishawant Kumar

Jawala Prakash

Jaydeep Dixit

Khushboo Priya

Madhulika G

Mahesh Sarade

Maitry

Manjunatha A

Mithun S

Mukesh

Muskan Garg

Nandani Gupta

Nishtha Jain

Nitin M

Prabir Kumar Satapathy

Prateek _iot

Prerna Singh

Rishav Dash

Saikumarreddy N

Sanjeev Kumar

Saurabh Shukla

Shiva Srivastava

Shivan K

Shivananda Sonwane

Shubham Sharma

Sowmiya Sivakumar

Sudhanshu Kumar

Suraj S Bilgi

Swati

Tarun

Wasim

Zeeshan

Time taken: 60.25 seconds, Fetched: 49 row(s)

hive> select distinct agent_name from agent_performance;

Query ID = cloudera_20221101210606_286d5cfa-e577-47db-9828-045875c5505f

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_1667316033891_0003, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667316033891_0003/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667316033891_0003

Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1

2022-11-01 21:06:19,465 Stage-1 map = 0%, reduce = 0%

2022-11-01 21:06:32,624 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.81 sec

2022-11-01 21:06:48,642 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6.63 sec

MapReduce Total cumulative CPU time: 6 seconds 630 msec

Ended Job = job_1667316033891_0003

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 6.63 sec   HDFS Read: 124583 HDFS Write: 841 SUCCESS

Total MapReduce CPU Time Spent: 6 seconds 630 msec

OK

Abhishek

Aditya

Aditya Shinde

Aditya_iot

Amersh

Ameya Jain

Anirudh

Ankit Sharma

Ankitjha

Anurag Tiwari

Aravind

Ashad Nasim

Ashish

Ayushi Mishra

Bharath

Boktiar Ahmed Bappy

Chaitra K Hiremath

Deepranjan Gupta

Dibyanshu

Harikrishnan Shaji

Hitesh Choudhary

Hrisikesh Neogi

Hyder Abbas

Ineuron Intelligence

Ishawant Kumar

Jawala Prakash

Jayant Kumar

Jaydeep Dixit

Khushboo Priya

Madhulika G

Mahak

Mahesh Sarade

Maitry

Maneesh

Manjunatha A

Mithun S

Mukesh

Mukesh Rao

Muskan Garg

Nandani Gupta

Nishtha Jain

Nitin M

Prabir Kumar Satapathy

Prateek _iot

Prerna Singh

Rishav Dash

Rohan

Saif Khan

Saikumarreddy N

Samprit

Sandipan Saha

Sanjeev Kumar

Sanjeevan

Saurabh Shukla

Shiva Srivastava

Shivan K

Shivan_S

Shivananda Sonwane

Shubham Sharma

Sowmiya Sivakumar

Spuri

Sudhanshu Kumar

Suraj S Bilgi

Swati

Tarun

Uday Mishra

Vasanth P

Vivek

Wasim

Zeeshan

Time taken: 46.158 seconds, Fetched: 71 row(s)

Q4) Find out agent average rating.

Solution:

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

hive> select agent_name as agent, avg(average_rating) as average_rating from agent_performance group by agent_name;

Query ID = cloudera_20221101211111_d3477195-298c-4632-97c5-05e0bc981a44

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_1667316033891_0006, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667316033891_0006/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667316033891_0006

Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1

2022-11-01 21:11:37,012 Stage-1 map = 0%, reduce = 0%

2022-11-01 21:11:44,688 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.33 sec

2022-11-01 21:11:57,761 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 4.79 sec

MapReduce Total cumulative CPU time: 4 seconds 790 msec

Ended Job = job_1667316033891_0006

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 4.79 sec   HDFS Read: 125817 HDFS Write: 1881 SUCCESS

Total MapReduce CPU Time Spent: 4 seconds 790 msec

OK

agent   average_rating

Abhishek   0.0

Aditya 0.0

Aditya Shinde   1.8003333409627278

Aditya_iot   2.3453333377838135

Amersh 0.0

Ameya Jain   2.21966667175293

Anirudh 0.6449999968210857

Ankit Sharma   0.0

Ankitjha   0.26666666666666666

Anurag Tiwari   0.18333333333333332

Aravind 2.1813333511352537

Ashad Nasim   0.16666666666666666

Ashish 0.0

Ayushi Mishra   3.481999969482422

Bharath 2.9836666584014893

Boktiar Ahmed Bappy   3.567999982833862

Chaitra K Hiremath   0.8646666606267294

Deepranjan Gupta   2.886666695276896

Dibyanshu   0.0

Harikrishnan Shaji   2.6396666526794434

Hitesh Choudhary   0.0

Hrisikesh Neogi 3.1363333304723104

Hyder Abbas   0.0

Ineuron Intelligence   0.0

Ishawant Kumar 3.543333347638448

Jawala Prakash 3.472000018755595

Jayant Kumar   1.068666664759318

Jaydeep Dixit   3.1670000314712525

Khushboo Priya 3.703666663169861

Madhulika G   3.4986666520436605

Mahak   0.1

Mahesh Sarade   2.4003333330154417

Maitry 2.9270000139872234

Maneesh 0.16666666666666666

Manjunatha A   3.5946666876475017

Mithun S   2.359000023206075

Mukesh 0.3096666653951009

Mukesh Rao   0.25566666523615517

Muskan Garg   0.712333329518636

Nandani Gupta   2.9236666679382326

Nishtha Jain   3.282333334287008

Nitin M 0.0

Prabir Kumar Satapathy 2.5103333314259846

Prateek _iot   2.4383333206176756

Prerna Singh   3.2326666434605915

Rishav Dash   1.4268333355585734

Rohan   0.0

Saif Khan   0.0

Saikumarreddy N 1.9803333441416422

Samprit 0.0

Sandipan Saha   0.4289999961853027

Sanjeev Kumar   3.3830000241597493

Sanjeevan   0.0

Saurabh Shukla 0.5556666692097981

Shiva Srivastava   0.9446666717529297

Shivan K   2.841333341598511

Shivan_S   0.14166666666666666

Shivananda Sonwane   4.232666659355163

Shubham Sharma 3.2253333568572997

Sowmiya Sivakumar   1.2599999984105428

Spuri   0.0

Sudhanshu Kumar 0.3333333333333333

Suraj S Bilgi   0.31200000445048015

Swati   2.4236666917800904

Tarun   0.05

Uday Mishra   0.0

Vasanth P   0.0

Vivek   0.5006666660308838

Wasim   2.400000015894572

Zeeshan 2.286999988555908

Time taken: 33.155 seconds, Fetched: 71 row(s)

Q5) Total working days for each agent

Solution:

hive> select agent_name as agent, count(distinct date) as number_of_working_days from agent_performance group by agent_name;

Query ID = cloudera_20221101211717_617f0d44-ffd7-46f9-b0a1-e8b032095df8

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_1667316033891_0008, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667316033891_0008/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667316033891_0008

Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1

2022-11-01 21:17:53,354 Stage-1 map = 0%, reduce = 0%

2022-11-01 21:18:01,959 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.62 sec

2022-11-01 21:18:13,514 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 5.15 sec

MapReduce Total cumulative CPU time: 5 seconds 150 msec

Ended Job = job_1667316033891_0008

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 5.15 sec   HDFS Read: 125596 HDFS Write: 1053 SUCCESS

Total MapReduce CPU Time Spent: 5 seconds 150 msec

OK

agent   number_of_working_days

       0

Abhishek   30

Aditya 30

Aditya Shinde   30

Aditya_iot   30

Amersh 30

Ameya Jain   30

Anirudh 30

Ankit Sharma   30

Ankitjha   30

Anurag Tiwari   30

Aravind 30

Ashad Nasim   30

Ashish 30

Ayushi Mishra   30

Bharath 30

Boktiar Ahmed Bappy   30

Chaitra K Hiremath   30

Deepranjan Gupta   30

Dibyanshu   30

Harikrishnan Shaji   30

Hitesh Choudhary   30

Hrisikesh Neogi 30

Hyder Abbas   30

Ineuron Intelligence   30

Ishawant Kumar 30

Jawala Prakash 30

Jayant Kumar   30

Jaydeep Dixit   30

Khushboo Priya 30

Madhulika G   30

Mahak   30

Mahesh Sarade   30

Maitry 30

Maneesh 30

Manjunatha A   30

Mithun S   30

Mukesh 30

Mukesh Rao   30

Muskan Garg   30

Nandani Gupta   30

Nishtha Jain   30

Nitin M 30

Prabir Kumar Satapathy 30

Prateek _iot   30

Prerna Singh   30

Rishav Dash   30

Rohan   30

Saif Khan   30

Saikumarreddy N 30

Samprit 30

Sandipan Saha   30

Sanjeev Kumar   30

Sanjeevan   30

Saurabh Shukla 30

Shiva Srivastava   30

Shivan K   30

Shivan_S   30

Shivananda Sonwane   30

Shubham Sharma 30

Sowmiya Sivakumar   30

Spuri   30

Sudhanshu Kumar 30

Suraj S Bilgi   30

Swati   30

Tarun   30

Uday Mishra   30

Vasanth P   30

Vivek   30

Wasim   30

Zeeshan 30

Time taken: 31.571 seconds, Fetched: 71 row(s)

Q6) Total query that each agent has taken

Solution:

hive> select agent_name as agent, sum(total_chats) as queries_taken from agent_performance group by agent_name;

Query ID = cloudera_20221101214848_5e4743f7-bfa2-437e-a8af-18d16fa84039

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_1667316033891_0009, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667316033891_0009/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667316033891_0009

Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1

2022-11-01 21:48:24,468 Stage-1 map = 0%, reduce = 0%

2022-11-01 21:48:33,352 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.17 sec

2022-11-01 21:48:44,544 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 4.69 sec

MapReduce Total cumulative CPU time: 4 seconds 690 msec

Ended Job = job_1667316033891_0009

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 4.69 sec   HDFS Read: 125339 HDFS Write: 1065 SUCCESS

Total MapReduce CPU Time Spent: 4 seconds 690 msec

OK

agent	queries_taken
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
Hitesh Choudhary	1
Hrisikesh Neogi	578
Hyder Abbas	0
Ineuron Intelligence	0
Ishawant Kumar	338
Jawala Prakash	439
Jayant Kumar	127
Jaydeep Dixit	512
Khushboo Priya	446
Madhulika G	469
Mahak	7
Mahesh Sarade	364
Maitry	542
Maneesh	4
Manjunatha A	413
Mithun S	503
Mukesh	19

Mukesh Rao   5

Muskan Garg   56

Nandani Gupta   560

Nishtha Jain   373

Nitin M 0

Prabir Kumar Satapathy   299

Prateek _iot   190

Prerna Singh   401

Rishav Dash   409

Rohan   0

Saif Khan   0

Saikumarreddy N 364

Samprit 1

Sandipan Saha   30

Sanjeev Kumar   507

Sanjeevan   0

Saurabh Shukla   16

Shiva Srivastava   53

Shivan K   357

Shivan_S   7

Shivananda Sonwane   441

Shubham Sharma   510

Sowmiya Sivakumar   206

Spuri   0

Sudhanshu Kumar 2

Suraj S Bilgi   28

Swati   524

Tarun   22

Uday Mishra   0

Vasanth P   0

Vivek   44

Wasim   433

Zeeshan 542

Time taken: 30.731 seconds, Fetched: 71 row(s)

Q7) Total Feedback that each agent has received

Solution:

hive> select agent_name as agent, sum(total_feedback) as feedbacks_received from agent_performance group by agent_name;

Query ID = cloudera_20221101215656_81013437-bff7-447c-b446-3709c703263a

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_1667316033891_0012, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667316033891_0012/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667316033891_0012

Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1

2022-11-01 21:56:17,193 Stage-1 map = 0%, reduce = 0%

2022-11-01 21:56:29,417 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.06 sec

2022-11-01 21:56:42,782 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6.62 sec

MapReduce Total cumulative CPU time: 6 seconds 620 msec

Ended Job = job_1667316033891_0012

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 6.62 sec   HDFS Read: 125346 HDFS Write: 1061 SUCCESS

Total MapReduce CPU Time Spent: 6 seconds 620 msec

OK

agent   feedbacks_received

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

Hitesh Choudhary   0

Hrisikesh Neogi 367

Hyder Abbas   0

Ineuron Intelligence   0

Ishawant Kumar 202

Jawala Prakash 250

Jayant Kumar   70

Jaydeep Dixit   305

Khushboo Priya 289

Madhulika G   281

Mahak   5

Mahesh Sarade   216

Maitry 347

Maneesh 3

Manjunatha A   254

Mithun S   364

Mukesh 17

Mukesh Rao   5

Muskan Garg   37

Nandani Gupta   308

Nishtha Jain   257

Nitin M 0

Prabir Kumar Satapathy 222

Prateek _iot   107

Prerna Singh   235

Rishav Dash   264

Rohan   0

Saif Khan   0

Saikumarreddy N 290

Samprit 0

Sandipan Saha   18

Sanjeev Kumar   311

Sanjeevan   0

Saurabh Shukla   8

Shiva Srivastava   46

Shivan K   243

Shivan_S   4

Shivananda Sonwane   263

Shubham Sharma   300

Sowmiya Sivakumar   141

Spuri   0

Sudhanshu Kumar   2

Suraj S Bilgi   15

Swati   302

Tarun   6

Uday Mishra   0

Vasanth P   0

Vivek   20

Wasim   284

Zeeshan   335

Time taken: 42.047 seconds, Fetched: 71 row(s)

Q8) Agent name who have average rating between 3.5 to 4

Solution:

hive> select agent_name as agent, avg(average_rating) as average_rating from agent_performance group by agent_name having average_rating between 3.5 and 4;

Query ID = cloudera_20221101220202_7dfde0dd-a9b0-4135-ab59-f4fa72fc5b7d

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_1667316033891_0013, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667316033891_0013/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667316033891_0013

Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1

2022-11-01 22:02:38,620 Stage-1 map = 0%, reduce = 0%

2022-11-01 22:02:48,489 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.04 sec

2022-11-01 22:02:59,284 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 4.85 sec

MapReduce Total cumulative CPU time: 4 seconds 850 msec

Ended Job = job_1667316033891_0013

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 4.85 sec   HDFS Read: 126449 HDFS Write: 136 SUCCESS

Total MapReduce CPU Time Spent: 4 seconds 850 msec

OK

agent   average_rating

Boktiar Ahmed Bappy   3.567999982833862

Ishawant Kumar 3.543333347638448

Khushboo Priya 3.703666663169861

Manjunatha A   3.5946666876475017

Time taken: 33.547 seconds, Fetched: 4 row(s)

Q9) Agent name who have rating less than 3.5

Solution:

hive> select agent_name as agent, avg(average_rating) as average_rating from agent_performance group by agent_name having average_rating < 3.5;

Query ID = cloudera_20221101220404_673bfe50-feae-4738-b904-6f241f2aa822

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_1667316033891_0014, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667316033891_0014/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667316033891_0014

Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1

2022-11-01 22:04:34,295 Stage-1 map = 0%, reduce = 0%

2022-11-01 22:04:46,500 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.71 sec

2022-11-01 22:05:03,498 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 7.32 sec

MapReduce Total cumulative CPU time: 7 seconds 320 msec

Ended Job = job_1667316033891_0014

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 7.32 sec   HDFS Read: 126252 HDFS Write: 1704 SUCCESS

Total MapReduce CPU Time Spent: 7 seconds 320 msec

OK

agent   average_rating

Abhishek   0.0

Aditya 0.0

Aditya Shinde   1.8003333409627278

Aditya_iot   2.3453333377838135

Amersh 0.0

Ameya Jain   2.21966667175293

Anirudh 0.6449999968210857

Ankit Sharma   0.0

Ankitjha   0.26666666666666666

Anurag Tiwari   0.18333333333333332

Aravind 2.1813333511352537

Ashad Nasim   0.16666666666666666

Ashish 0.0

Ayushi Mishra   3.481999969482422

Bharath 2.9836666584014893

Chaitra K Hiremath   0.8646666606267294

Deepranjan Gupta   2.886666695276896

Dibyanshu   0.0

Harikrishnan Shaji   2.6396666526794434

Hitesh Choudhary   0.0

Hrisikesh Neogi 3.1363333304723104

Hyder Abbas   0.0

Ineuron Intelligence   0.0

Jawala Prakash 3.472000018755595

Jayant Kumar   1.068666664759318

Jaydeep Dixit   3.1670000314712525

Madhulika G   3.4986666520436605

Mahak   0.1

Mahesh Sarade   2.4003333330154417

Maitry 2.9270000139872234

Maneesh 0.16666666666666666

Mithun S   2.359000023206075

Mukesh 0.3096666653951009

Mukesh Rao   0.25566666523615517

Muskan Garg   0.712333329518636

Nandani Gupta   2.9236666679382326

Nishtha Jain   3.282333334287008

Nitin M 0.0

Prabir Kumar Satapathy 2.5103333314259846

Prateek _iot   2.4383333206176756

Prerna Singh   3.2326666434605915

Rishav Dash   1.4268333355585734

Rohan   0.0

Saif Khan   0.0

Saikumarreddy N 1.9803333441416422

Samprit 0.0

Sandipan Saha   0.4289999961853027

Sanjeev Kumar   3.3830000241597493

Sanjeevan   0.0

Saurabh Shukla 0.5556666692097981

Shiva Srivastava   0.9446666717529297

Shivan K   2.841333341598511

Shivan_S   0.14166666666666666

Shubham Sharma 3.2253333568572997

Sowmiya Sivakumar   1.2599999984105428

Spuri   0.0

Sudhanshu Kumar 0.3333333333333333

Suraj S Bilgi   0.31200000445048015

Swati   2.4236666917800904

Tarun   0.05

Uday Mishra   0.0

Vasanth P   0.0

Vivek   0.5006666660308838

Wasim   2.400000015894572

Zeeshan 2.286999988555908

Time taken: 45.33 seconds, Fetched: 65 row(s)

Q10) Agent name who have rating more than 4.5

Solution:

hive> select agent_name as agent, avg(average_rating) as average_rating from agent_performance group by agent_name having average_rating > 4.5;

Query ID = cloudera_20221101220505_cc78d0b4-99be-4cd0-8227-fe4d79e36f40

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_1667316033891_0015, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667316033891_0015/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667316033891_0015

Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1

2022-11-01 22:05:30,465 Stage-1 map = 0%, reduce = 0%

2022-11-01 22:05:43,975 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.94 sec

2022-11-01 22:05:59,236 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6.62 sec

MapReduce Total cumulative CPU time: 6 seconds 620 msec

Ended Job = job_1667316033891_0015

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 6.62 sec   HDFS Read: 126251 HDFS Write: 0 SUCCESS

Total MapReduce CPU Time Spent: 6 seconds 620 msec

OK

agent   average_rating

Time taken: 44.717 seconds

Q11) How many feedback agents have received more than 4.5 average

Solution:

hive> select count(*) from(select agent_name as agent, avg(total_feedback) as average_feedback from agent_performance group by agent_name having average_feedback > 4.5)t;

Query ID = cloudera_20221101221919_c687ebc1-ccb0-4d82-bd3b-2bb333af0f74

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_1667316033891_0019, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667316033891_0019/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667316033891_0019

Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1

2022-11-01 22:20:02,892 Stage-1 map = 0%, reduce = 0%

2022-11-01 22:20:14,116 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.71 sec

2022-11-01 22:20:27,703 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6.64 sec

MapReduce Total cumulative CPU time: 6 seconds 640 msec

Ended Job = job_1667316033891_0019

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_1667316033891_0020, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667316033891_0020/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667316033891_0020

Hadoop job information for Stage-2: number of mappers: 1; number of reducers: 1

2022-11-01 22:20:42,088 Stage-2 map = 0%, reduce = 0%

2022-11-01 22:20:52,198 Stage-2 map = 100%, reduce = 0%, Cumulative CPU 2.39 sec

2022-11-01 22:21:04,433 Stage-2 map = 100%, reduce = 100%, Cumulative CPU 5.45 sec

MapReduce Total cumulative CPU time: 5 seconds 450 msec

Ended Job = job_1667316033891_0020

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 6.64 sec   HDFS Read: 125695 HDFS Write: 114 SUCCESS

Stage-Stage-2: Map: 1 Reduce: 1   Cumulative CPU: 5.45 sec   HDFS Read: 4550 HDFS Write: 3 SUCCESS

Total MapReduce CPU Time Spent: 12 seconds 90 msec

OK

_c0

31

Time taken: 77.471 seconds, Fetched: 1 row(s)

Q12) average weekly response time for each agent

Solution:

hive> select agent, avg(weekly_response_time_in_sec) as avg_weekly_response_time_in_sec from (select week, agent, sum((time[0]*3600+time[1]*60+time[2])) as weekly_response_time_in_sec from(select agent_name as agent, weekofyear(date) as week, split(average_response_time,':') as time from agent_performance) t group by agent, week)s group by agent;

Query ID = cloudera_20221101233535_4ba561a1-9387-4cca-a257-2a3de2da2ae0

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_1667316033891_0027, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667316033891_0027/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667316033891_0027

Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1

2022-11-01 23:35:17,893 Stage-1 map = 0%, reduce = 0%

2022-11-01 23:35:27,189 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.21 sec

2022-11-01 23:35:37,351 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 5.32 sec

MapReduce Total cumulative CPU time: 5 seconds 320 msec

Ended Job = job_1667316033891_0027

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 5.32 sec   HDFS Read: 129398 HDFS Write: 1212 SUCCESS

Total MapReduce CPU Time Spent: 5 seconds 320 msec

OK

agent   avg_weekly_response_time_in_sec

Abhishek   0.0

Aditya 0.0

Aditya Shinde   178.4

Aditya_iot   203.8

Amersh 0.0

Ameya Jain   126.8

Anirudh 130.6

Ankit Sharma   0.0

Ankitjha   26.6

Anurag Tiwari   50.6

Aravind 128.2

Ashad Nasim   231.8

Ashish 0.0

Ayushi Mishra   362.0

Bharath 160.8

Boktiar Ahmed Bappy   396.2

Chaitra K Hiremath   90.8

Deepranjan Gupta   319.2

Dibyanshu   7.6

Harikrishnan Shaji   203.8

Hitesh Choudhary   0.0

Hrisikesh Neogi 303.0

Hyder Abbas   0.0

Ineuron Intelligence   0.0

Ishawant Kumar 300.8

Jawala Prakash 565.4

Jayant Kumar   110.6

Jaydeep Dixit   266.4

Khushboo Priya 367.8

Madhulika G   398.6

Mahak   0.0

Mahesh Sarade   278.6

Maitry 383.0

Maneesh 27.0

Manjunatha A   217.0

Mithun S   173.6

Mukesh 20.0

Mukesh Rao   78.8

Muskan Garg   35.6

Nandani Gupta   359.2

Nishtha Jain   364.8

Nitin M 0.0

Prabir Kumar Satapathy 228.0

Prateek _iot   135.0

Prerna Singh   286.0

Rishav Dash   363.8

Rohan   0.0

Saif Khan   0.0

Saikumarreddy N 151.0

Samprit 0.0

Sandipan Saha   35.4

Sanjeev Kumar   307.2

Sanjeevan   0.0

Saurabh Shukla 21.0

Shiva Srivastava   60.0

Shivan K   287.4

Shivan_S   14.6

Shivananda Sonwane   336.0

Shubham Sharma 290.0

Sowmiya Sivakumar   157.0

Spuri   0.0

Sudhanshu Kumar 24.0

Suraj S Bilgi   36.4

Swati   346.8

Tarun   0.0

Uday Mishra   0.0

Vasanth P   0.0

Vivek   82.2

Wasim   178.2

Zeeshan 370.4

Time taken: 30.063 seconds, Fetched: 71 row(s)

Q13) average weekly resolution time for each agents

Solution:

hive> select agent, avg(weekly_resolution_time_in_sec) as avg_weekly_resolution_time_in_sec from (select week, agent, sum((time[0]*3600+time[1]*60+time[2])) as weekly_resolution_time_in_sec from(select agent_name as agent, weekofyear(date) as week, split(average_resolution_time,':') as time from agent_performance) t group by agent, week)s group by agent;

Query ID = cloudera_20221101234040_71190205-b2a4-41ad-b17e-5be389b17f6a

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_1667316033891_0028, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667316033891_0028/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667316033891_0028

Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1

2022-11-01 23:40:16,339 Stage-1 map = 0%, reduce = 0%

2022-11-01 23:40:28,723 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 4.64 sec

2022-11-01 23:40:49,453 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 11.52 sec

MapReduce Total cumulative CPU time: 11 seconds 520 msec

Ended Job = job_1667316033891_0028

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 11.52 sec   HDFS Read: 129404 HDFS Write: 1271 SUCCESS

Total MapReduce CPU Time Spent: 11 seconds 520 msec

OK

agent   avg_weekly_resolution_time_in_sec

Abhishek   0.0

Aditya 0.0

Aditya Shinde   3723.8

Aditya_iot   3535.8

Amersh 0.0

Ameya Jain   1971.0

Anirudh 1112.8

Ankit Sharma   0.0

Ankitjha   326.8

Anurag Tiwari   443.6

Aravind 3213.4

Ashad Nasim   125.6

Ashish 0.0

Ayushi Mishra   5535.6

Bharath 3885.6

Boktiar Ahmed Bappy   6143.0

Chaitra K Hiremath   533.8

Deepranjan Gupta   7507.0

Dibyanshu   148.0

Harikrishnan Shaji   4070.6

Hitesh Choudhary   17.0

Hrisikesh Neogi 5557.8

Hyder Abbas   0.0

Ineuron Intelligence   0.0

Ishawant Kumar 5190.6

Jawala Prakash 4682.6

Jayant Kumar   1769.6

Jaydeep Dixit   6486.4

Khushboo Priya 6026.6

Madhulika G   5595.4

Mahak   240.8

Mahesh Sarade   3297.0

Maitry 4602.4

Maneesh 250.0

Manjunatha A   6105.2

Mithun S   2080.0

Mukesh 455.6

Mukesh Rao   2763.4

Muskan Garg   691.2

Nandani Gupta   6538.2

Nishtha Jain   3350.4

Nitin M 0.0

Prabir Kumar Satapathy 2139.8

Prateek _iot   2917.4

Prerna Singh   5948.4

Rishav Dash   6114.6

Rohan   0.0

Saif Khan   0.0

Saikumarreddy N 2240.2

Samprit 20.6

Sandipan Saha   947.0

Sanjeev Kumar   6189.8

Sanjeevan   0.0

Saurabh Shukla 428.4

Shiva Srivastava   532.8

Shivan K   5709.6

Shivan_S   220.4

Shivananda Sonwane   7611.6

Shubham Sharma 6259.0

Sowmiya Sivakumar   2144.0

Spuri   0.0

Sudhanshu Kumar 701.8

Suraj S Bilgi   946.0

Swati   4419.4

Tarun   542.8

Uday Mishra   0.0

Vasanth P   0.0

Vivek   916.6

Wasim   4133.2

Zeeshan 3870.4

Time taken: 48.741 seconds, Fetched: 71 row(s)

Q14) Find the number of chats on which they have received a feedback

Solution:

hive> select agent_name as agent, sum(total_chats) as chats, sum(total_feedback) as chats_with_feedback_received from agent_performance group by agent_name;

Query ID = cloudera_20221101234747_993bf134-206c-4ae0-9001-cba22ef8a2a8

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_1667316033891_0029, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667316033891_0029/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667316033891_0029

Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1

2022-11-01 23:47:53,761 Stage-1 map = 0%, reduce = 0%

2022-11-01 23:48:06,087 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.01 sec

2022-11-01 23:48:19,713 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6.5 sec

MapReduce Total cumulative CPU time: 6 seconds 500 msec

Ended Job = job_1667316033891_0029

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 6.5 sec   HDFS Read: 125793 HDFS Write: 1285 SUCCESS

Total MapReduce CPU Time Spent: 6 seconds 500 msec

OK

agent	chats	chats_with_feedback_received
Abhishek	0	0
Aditya	0	0
Aditya Shinde	277	153
Aditya_iot	231	131
Amersh	0	0
Ameya Jain	322	228
Anirudh	81	39
Ankit Sharma	0	0
Ankitjha	5	3
Anurag Tiwari	4	3
Aravind	366	233
Ashad Nasim	18	9
Ashish	0	0
Ayushi Mishra	514	329
Bharath	369	247
Boktiar Ahmed Bappy	452	311
Chaitra K Hiremath	64	37
Deepranjan Gupta	493	312
Dibyanshu	1	0
Harikrishnan Shaji	381	231
Hitesh Choudhary	1	0
Hrisikesh Neogi	578	367
Hyder Abbas	0	0
Ineuron Intelligence	0	0
Ishawant Kumar	338	202
Jawala Prakash	439	250
Jayant Kumar	127	70
Jaydeep Dixit	512	305
Khushboo Priya	446	289
Madhulika G	469	281
Mahak	7	5
Mahesh Sarade	364	216
Maitry	542	347
Maneesh	4	3

Manjunatha A   413   254

Mithun S   503   364

Mukesh 19   17

Mukesh Rao   5   5

Muskan Garg   56   37

Nandani Gupta   560   308

Nishtha Jain   373   257

Nitin M 0   0

Prabir Kumar Satapathy 299   222

Prateek _iot   190   107

Prerna Singh   401   235

Rishav Dash   409   264

Rohan   0   0

Saif Khan   0   0

Saikumarreddy N 364   290

Samprit 1   0

Sandipan Saha   30   18

Sanjeev Kumar   507   311

Sanjeevan   0   0

Saurabh Shukla 16   8

Shiva Srivastava   53   46

Shivan K   357   243

Shivan_S   7   4

Shivananda Sonwane   441   263

Shubham Sharma 510   300

Sowmiya Sivakumar   206   141

Spuri   0   0

Sudhanshu Kumar 2   2

Suraj S Bilgi   28   15

Swati   524   302

Tarun   22   6

Uday Mishra   0   0

Vasanth P   0   0

Vivek   44   20

Wasim   433   284

Zeeshan 542   335

Time taken: 41.182 seconds, Fetched: 71 row(s)

Q15) Total contribution hour for each and every agents weekly basis

Solution:

hive> select week, agent, sum((time[0]*3600+time[1]*60+time[2])/3600) as total_hrs_contributed from(select agent, weekofyear(date) as week, split(duration,':') as time from agent_loging) t group by agent, week;

Query ID = cloudera_20221102000303_0195e12e-5cf9-4ba2-9b71-195b9943d4ee

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_1667316033891_0032, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667316033891_0032/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667316033891_0032

Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1

2022-11-02 00:03:43,956 Stage-1 map = 0%, reduce = 0%

2022-11-02 00:03:58,926 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 4.5 sec

2022-11-02 00:04:13,594 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 9.45 sec

MapReduce Total cumulative CPU time: 9 seconds 450 msec

Ended Job = job_1667316033891_0032

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 9.45 sec HDFS Read: 68662 HDFS Write: 3068 SUCCESS

Total MapReduce CPU Time Spent: 9 seconds 450 msec

OK

week	agent	total_hrs_contributed
30	Aditya Shinde	0.03611111111111111
29	Aditya_iot	6.095277777777778
30	Aditya_iot	9.635833333333334
30	Amersh	3.0638888888888887
29	Ameya Jain	24.083055555555553
30	Ameya Jain	17.9925
30	Ankitjha	2.2669444444444444
29	Anurag Tiwari	0.2644444444444444

30	Anurag Tiwari	2.5144444444444445
29	Aravind	24.23555555555555
30	Aravind	0.06361111111111112
29	Ayushi Mishra	17.790277777777778
30	Ayushi Mishra	20.33138888888889
29	Bharath	24.070833333333333
30	Bharath	24.005833333333335
29	Boktiar Ahmed Bappy	17.75027777777778
30	Boktiar Ahmed Bappy	22.518333333333334
29	Chaitra K Hiremath	2.2347222222222225
30	Chaitra K Hiremath	32.090833333333336
29	Deepranjan Gupta	48.99638888888889
30	Deepranjan Gupta	57.27888888888887
29	Dibyanshu	27.743888888888907
30	Dibyanshu	24.851944444444474
29	Harikrishnan Shaji	21.438333333333333
30	Harikrishnan Shaji	32.27638888888889
29	Hrisikesh Neogi	26.89138888888889
30	Hrisikesh Neogi	30.677222222222223
29	Hyder Abbas	0.33611111111111114
30	Hyder Abbas	0.051944444444444446
29	Ineuron Intelligence	1.448611111111111
29	Ishawant Kumar	25.72083333333333
30	Ishawant Kumar	26.05805555555555
29	Jawala Prakash	24.340000000000003
30	Jawala Prakash	22.062222222222225
29	Jaydeep Dixit	41.91444444444444
30	Jaydeep Dixit	17.926111111111112
29	Khushboo Priya	21.715833333333336
30	Khushboo Priya	21.84277777777778
29	Madhulika G	25.850833333333334
30	Madhulika G	20.209444444444447
29	Mahesh Sarade	25.483055555555552
30	Mahesh Sarade	17.438888888888886
29	Maitry	24.65777777777778
30	Maitry	6.287222222222222
29	Manjunatha A	18.351111111111113
30	Manjunatha A	22.92361111111111

29	Mithun S	17.37972222222222
30	Mithun S	27.79388888888889
30	Mukesh	8.905
29	Muskan Garg	3.318611111111111
30	Muskan Garg	14.016944444444446
29	Nandani Gupta	17.33388888888889
30	Nandani Gupta	22.83833333333333
29	Nishtha Jain	22.11583333333333
30	Nishtha Jain	21.736944444444443
29	Nitin M	0.7988888888888889
29	Prabir Kumar Satapathy	17.524444444444445
30	Prabir Kumar Satapathy	15.852777777777776
29	Prateek _iot	7.269722222222223
30	Prateek _iot	11.148333333333335
29	Prerna Singh	18.517500000000002
30	Prerna Singh	27.19805555555556
29	Rishav Dash	18.89388888888889
30	Rishav Dash	22.881666666666668
29	Saikumarreddy N	24.980555555555558
30	Saikumarreddy N	18.156944444444445
29	Sanjeev Kumar	19.360833333333332
30	Sanjeev Kumar	25.326388888888893
29	Saurabh Shukla	16.663055555555555
29	Shiva Srivastava	1.906111111111111
30	Shiva Srivastava	13.088055555555556
29	Shivan K	16.71388888888889
30	Shivan K	19.388333333333332
29	Shivananda Sonwane	20.83416666666667
30	Shivananda Sonwane	28.45361111111111
29	Shubham Sharma	30.510277777777784
30	Shubham Sharma	23.28805555555555
29	Sowmiya Sivakumar	17.065833333333334
30	Sowmiya Sivakumar	27.688333333333333
29	Sudhanshu Kumar	24.454722222222223
30	Sudhanshu Kumar	21.77638888888889
30	Suraj S Bilgi	12.559166666666666
29	Swati	18.85861111111111
30	Swati	6.1425

26   Tarun   10.13888888888889

29   Wasim   19.625555555555554

30   Wasim   28.553611111111113

29   Zeeshan 24.427500000000002

30   Zeeshan 24.661111111111115

Time taken: 45.08 seconds, Fetched: 89 row(s)

Q16) 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.

Solution:

Inner Join:

hive> select l.*, p.*

   > from

   > agent_loging l

   > inner join

   > agent_performance p

   > on l.agent = p.agent_name

   > limit 5;

Query ID = cloudera_20221102012525_d501b321-d4d3-421b-8138-c772966267be

Total jobs = 1

Execution log at: /tmp/cloudera/cloudera_20221102012525_d501b321-d4d3-421b-8138-c772966267be.log

2022-11-02 01:25:37   Starting to launch local task to process map join;   maximum memory = 932184064

2022-11-02 01:25:44   Dump the side-table for tag: 0 with group count: 49 into file: file:/tmp/cloudera/a7e5d834-a1dc-4221-bc58-c45c9866afe5/hive_2022-11-02_01-25-18_397_2973530280533039045-1/-local-10003/HashTable-Stage-3/MapJoin-mapfile00--.hashtable

2022-11-02 01:25:44   Uploaded 1 File to: file:/tmp/cloudera/a7e5d834-a1dc-4221-bc58-c45c9866afe5/hive_2022-11-02_01-25-18_397_2973530280533039045-1/-local-10003/HashTable-Stage-3/MapJoin-mapfile00--.hashtable (39341 bytes)

2022-11-02 01:25:44   End of local task; Time Taken: 6.518 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_1667377030429_0001, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667377030429_0001/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667377030429_0001

Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0

2022-11-02 01:26:19,004 Stage-3 map = 0%, reduce = 0%

2022-11-02 01:26:37,823 Stage-3 map = 100%, reduce = 0%, Cumulative CPU 4.27 sec

MapReduce Total cumulative CPU time: 4 seconds 270 msec

Ended Job = job_1667377030429_0001

MapReduce Jobs Launched:

Stage-Stage-3: Map: 1   Cumulative CPU: 4.27 sec   HDFS Read: 13194 HDFS Write: 542 SUCCESS

Total MapReduce CPU Time Spent: 4 seconds 270 msec

OK

l.s_no	l.agent	l.date	l.login_time	l.logout_time	l.duration	p.s_no	p.date	p.agent_name	p.total_chats	p.average_response_time	p.average_resolution_time	p.average_rating	p.total_feedback
16	Prerna Singh	2022-07-30	12:32:28	14:10:08	01:37:40	1	2022-07-30	Prerna Singh	11	00:00:38	00:04:20	4.11	9
75	Prerna Singh	2022-07-29	17:47:06	21:03:44	03:16:37	1	2022-07-30	Prerna Singh	11	00:00:38	00:04:20	4.11	9
91	Prerna Singh	2022-07-29	15:08:22	17:20:49	02:12:27	1	2022-07-30	Prerna Singh	11	00:00:38	00:04:20	4.11	9
110	Prerna Singh	2022-07-29	12:08:23	12:11:35	00:03:11	1	2022-07-30	Prerna Singh	11	00:00:38	00:04:20	4.11	9
336	Prerna Singh	2022-07-27	13:11:06	20:58:35	07:47:29	1	2022-07-30	Prerna Singh	11	00:00:38	00:04:20	4.11	9

Time taken: 81.761 seconds, Fetched: 5 row(s)

Left Join:

hive> select l.*, p.*

    > from

    > agent_loging l

    > left join

    > agent_performance p

    > on l.agent = p.agent_name

    > limit 5;

Query ID = cloudera_20221102012727_6495894f-c7bf-444f-a20f-ab60098d710d

Total jobs = 1

Execution log at: /tmp/cloudera/cloudera_20221102012727_6495894f-c7bf-444f-a20f-ab60098d710d.log

2022-11-02 01:27:50   Starting to launch local task to process map join;   maximum memory = 932184064

2022-11-02 01:27:52   Dump the side-table for tag: 1 with group count: 71 into file: file:/tmp/cloudera/a7e5d834-a1dc-4221-bc58-c45c9866afe5/hive_2022-11-02_01-27-41_896_8829501887617462410-1/-local-10003/HashTable-Stage-3/MapJoin-mapfile11--.hashtable

2022-11-02 01:27:53   Uploaded 1 File to: file:/tmp/cloudera/a7e5d834-a1dc-4221-bc58-c45c9866afe5/hive_2022-11-02_01-27-41_896_8829501887617462410-1/-local-10003/HashTable-Stage-3/MapJoin-mapfile11--.hashtable (77631 bytes)

2022-11-02 01:27:53   End of local task; Time Taken: 3.103 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_1667377030429_0002, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667377030429_0002/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667377030429_0002

Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0

2022-11-02 01:28:10,612 Stage-3 map = 0%, reduce = 0%

2022-11-02 01:28:28,294 Stage-3 map = 100%, reduce = 0%, Cumulative CPU 4.95 sec

MapReduce Total cumulative CPU time: 4 seconds 950 msec

Ended Job = job_1667377030429_0002

MapReduce Jobs Launched:

Stage-Stage-3: Map: 1   Cumulative CPU: 4.95 sec   HDFS Read: 13247 HDFS Write: 600 SUCCESS

Total MapReduce CPU Time Spent: 4 seconds 950 msec

OK

l.s_no	l.agent	l.date	l.login_time	l.logout_time	l.duration	p.s_no	p.date	p.agent_name	p.total_chats	p.average_response_time	p.average_resolution_time	p.average_rating	p.total_feedback
1	Shivananda Sonwane	2022-07-30	15:35:29	17:39:39	02:04:10	69	2022-07-30	Shivananda Sonwane	4	00:01:14	00:16:53	5.0	1
1	Shivananda Sonwane	2022-07-30	15:35:29	17:39:39	02:04:10	73	2022-07-29	Shivananda Sonwane	14	00:00:45	00:15:38	4.67	9
1	Shivananda Sonwane	2022-07-30	15:35:29	17:39:39	02:04:10	214	2022-07-28	Shivananda Sonwane	5	00:00:31	00:38:04	5.0	4
1	Shivananda Sonwane	2022-07-30	15:35:29	17:39:39	02:04:10	285	2022-07-27	Shivananda Sonwane	26	00:01:12	00:20:10	4.22	18
1	Shivananda Sonwane	2022-07-30	15:35:29	17:39:39	02:04:10	360	2022-07-26	Shivananda Sonwane	24	00:00:51	00:22:28	5.0	14

Time taken: 47.602 seconds, Fetched: 5 row(s)

Right Join:

hive> select l.*, p.*

    > from

    > agent_loging l

    > right join

    > agent_performance p

    > on l.agent = p.agent_name

    > limit 3;

Query ID = cloudera_20221102013131_5a7932b2-a989-4ece-8444-63b3a7a724b6

Total jobs = 1

Execution log at: /tmp/cloudera/cloudera_20221102013131_5a7932b2-a989-4ece-8444-63b3a7a724b6.log

2022-11-02 01:31:28   Starting to launch local task to process map join;   maximum memory = 932184064

2022-11-02 01:31:31   Dump the side-table for tag: 0 with group count: 49 into file: file:/tmp/cloudera/a7e5d834-a1dc-4221-bc58-c45c9866afe5/hive_2022-11-02_01-31-18_589_6374086923008905591-1/-local-10003/HashTable-Stage-3/MapJoin-mapfile20--.hashtable

2022-11-02 01:31:31   Uploaded 1 File to: file:/tmp/cloudera/a7e5d834-a1dc-4221-bc58-c45c9866afe5/hive_2022-11-02_01-31-18_589_6374086923008905591-1/-local-10003/HashTable-Stage-3/MapJoin-mapfile20--.hashtable (39341 bytes)

2022-11-02 01:31:31   End of local task; Time Taken: 2.557 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_1667377030429_0003, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667377030429_0003/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667377030429_0003

Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0

2022-11-02 01:31:46,028 Stage-3 map = 0%, reduce = 0%

2022-11-02 01:32:03,722 Stage-3 map = 100%, reduce = 0%, Cumulative CPU 3.86 sec

MapReduce Total cumulative CPU time: 3 seconds 860 msec

Ended Job = job_1667377030429_0003

MapReduce Jobs Launched:

Stage-Stage-3: Map: 1   Cumulative CPU: 3.86 sec   HDFS Read: 13221 HDFS Write: 324 SUCCESS

Total MapReduce CPU Time Spent: 3 seconds 860 msec

OK

l.s_no	l.agent	l.date	l.login_time	l.logout_time	l.duration	p.s_no	p.date	p.agent_name	p.total_chats	p.average_response_time	p.average_resolution_time	p.average_rating	p.total_feedback
16	Prerna Singh	2022-07-30	12:32:28	14:10:08	01:37:40	1	2022-07-30	Prerna Singh	11	00:00:38	00:04:20	4.11	9
75	Prerna Singh	2022-07-29	17:47:06	21:03:44	03:16:37	1	2022-07-30	Prerna Singh	11	00:00:38	00:04:20	4.11	9
91	Prerna Singh	2022-07-29	15:08:22	17:20:49	02:12:27	1	2022-07-30	Prerna Singh	11	00:00:38	00:04:20	4.11	9

Time taken: 46.342 seconds, Fetched: 3 row(s)

Export data into local system:

Inner Join:

[cloudera@quickstart ~]$ hive -e 'select l.*, p.* from agent.agent_loging l inner join agent.agent_performance p on l.agent = p.agent_name limit 5' > /tmp/Agent_data/agent_inner_join.csv

Logging initialized using configuration in file:/etc/hive/conf.dist/hive-log4j.properties

Query ID = cloudera_20221102021313_3dbb874d-fdc2-49bd-b1d6-f3eb47b5c6d5

Total jobs = 1

Execution log at: /tmp/cloudera/cloudera_20221102021313_3dbb874d-fdc2-49bd-b1d6-f3eb47b5c6d5.log

2022-11-02 02:13:23 Starting to launch local task to process map join;   maximum memory = 932184064

2022-11-02 02:13:27 Dump the side-table for tag: 0 with group count: 49 into file: file:/tmp/cloudera/53d10f02-679f-4b61-918e-75477428efc2/hive_2022-11-02_02-13-11_168_5834537046314652500-1/-local-10003/HashTable-Stage-3/MapJoin-mapfile00--.hashtable

2022-11-02 02:13:27 Uploaded 1 File to: file:/tmp/cloudera/53d10f02-679f-4b61-918e-75477428efc2/hive_2022-11-02_02-13-11_168_5834537046314652500-1/-local-10003/HashTable-Stage-3/MapJoin-mapfile00--.hashtable (39341 bytes)

2022-11-02 02:13:27 End of local task; Time Taken: 4.238 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_1667377030429_0006, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667377030429_0006/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667377030429_0006

Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0

2022-11-02 02:13:47,531 Stage-3 map = 0%, reduce = 0%

2022-11-02 02:14:05,222 Stage-3 map = 100%, reduce = 0%, Cumulative CPU 5.09 sec

MapReduce Total cumulative CPU time: 5 seconds 90 msec

Ended Job = job_1667377030429_0006

MapReduce Jobs Launched:

Stage-Stage-3: Map: 1   Cumulative CPU: 5.09 sec   HDFS Read: 13194 HDFS Write: 542 SUCCESS

Total MapReduce CPU Time Spent: 5 seconds 90 msec

OK

Time taken: 56.553 seconds, Fetched: 5 row(s)

Left Join:

[cloudera@quickstart ~]$ hive -e 'select l.*, p.* from agent.agent_loging l left join agent.agent_performance p on l.agent = p.agent_name limit 5' > /tmp/Agent_data/agent_left_join.csv

Logging initialized using configuration in file:/etc/hive/conf.dist/hive-log4j.properties

Query ID = cloudera_20221102021414_f7ab8dd8-fdc3-448a-a846-512bbd2451c6

Total jobs = 1

Execution log at: /tmp/cloudera/cloudera_20221102021414_f7ab8dd8-fdc3-448a-a846-512bbd2451c6.log

2022-11-02 02:14:54 Starting to launch local task to process map join;   maximum memory = 932184064

2022-11-02 02:14:57 Dump the side-table for tag: 1 with group count: 71 into file: file:/tmp/cloudera/4cf935ee-290a-4387-bccc-1d023b332b07/hive_2022-11-02_02-14-43_629_6565383743414948191-1/-local-10003/HashTable-Stage-3/MapJoin-mapfile01--.hashtable

2022-11-02 02:14:57 Uploaded 1 File to: file:/tmp/cloudera/4cf935ee-290a-4387-bccc-1d023b332b07/hive_2022-11-02_02-14-43_629_6565383743414948191-1/-local-10003/HashTable-Stage-3/MapJoin-mapfile01--.hashtable (77631 bytes)

2022-11-02 02:14:57 End of local task; Time Taken: 3.059 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_1667377030429_0007, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667377030429_0007/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667377030429_0007

Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0

2022-11-02 02:15:14,314 Stage-3 map = 0%, reduce = 0%

2022-11-02 02:15:31,808 Stage-3 map = 100%, reduce = 0%, Cumulative CPU 4.5 sec

MapReduce Total cumulative CPU time: 4 seconds 500 msec

Ended Job = job_1667377030429_0007

MapReduce Jobs Launched:

Stage-Stage-3: Map: 1   Cumulative CPU: 4.5 sec   HDFS Read: 13055 HDFS Write: 600 SUCCESS

Total MapReduce CPU Time Spent: 4 seconds 500 msec

OK

Time taken: 50.644 seconds, Fetched: 5 row(s)

Right Join:

[cloudera@quickstart ~]$ hive -e 'select l.*, p.* from agent.agent_loging l right join agent.agent_performance p on l.agent = p.agent_name limit 5' > /tmp/Agent_data/agent_right_join.csv

Logging initialized using configuration in file:/etc/hive/conf.dist/hive-log4j.properties

Query ID = cloudera_20221102021616_6c0ea584-01f3-4d15-b707-8b2116b0307d

Total jobs = 1

Execution log at: /tmp/cloudera/cloudera_20221102021616_6c0ea584-01f3-4d15-b707-8b2116b0307d.log

2022-11-02 02:16:20 Starting to launch local task to process map join;   maximum memory = 932184064

2022-11-02 02:16:22 Dump the side-table for tag: 0 with group count: 49 into file: file:/tmp/cloudera/7ef817ad-963c-4d3b-b113-e87e225620e4/hive_2022-11-02_02-16-08_907_5363628296010565902-1/-local-10003/HashTable-Stage-3/MapJoin-mapfile00--.hashtable

2022-11-02 02:16:23 Uploaded 1 File to: file:/tmp/cloudera/7ef817ad-963c-4d3b-b113-e87e225620e4/hive_2022-11-02_02-16-08_907_5363628296010565902-1/-local-10003/HashTable-Stage-3/MapJoin-mapfile00--.hashtable (39341 bytes)

2022-11-02 02:16:23 End of local task; Time Taken: 3.068 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_1667377030429_0008, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667377030429_0008/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667377030429_0008

Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0

2022-11-02 02:16:44,458 Stage-3 map = 0%, reduce = 0%

2022-11-02 02:16:59,648 Stage-3 map = 100%, reduce = 0%, Cumulative CPU 4.45 sec

MapReduce Total cumulative CPU time: 4 seconds 450 msec

Ended Job = job_1667377030429_0008

MapReduce Jobs Launched:

Stage-Stage-3: Map: 1   Cumulative CPU: 4.45 sec   HDFS Read: 13029 HDFS Write: 542 SUCCESS

Total MapReduce CPU Time Spent: 4 seconds 450 msec

OK

Time taken: 53.231 seconds, Fetched: 5 row(s)

Check the exported data in local system:

[cloudera@quickstart ~]$ ls /tmp/Agent_data

agent_inner_join.csv agent_left_join.csv AgentLogingReport.csv AgentPerformance.csv agent_right_join.csv

[cloudera@quickstart ~]$ cat /tmp/Agent_data/agent_inner_join.csv

16	Prerna Singh	2022-07-30	12:32:28	14:10:08	01:37:40	1	2022-07-30	Prerna Singh	11	00:00:38	00:04:20
4.11	9										
75	Prerna Singh	2022-07-29	17:47:06	21:03:44	03:16:37	1	2022-07-30	Prerna Singh	11	00:00:38	00:04:20
4.11	9										
91	Prerna Singh	2022-07-29	15:08:22	17:20:49	02:12:27	1	2022-07-30	Prerna Singh	11	00:00:38	00:04:20
4.11	9										
110	Prerna Singh	2022-07-29	12:08:23	12:11:35	00:03:11	1	2022-07-30	Prerna Singh	11	00:00:38	00:04:20
4.11	9										
336	Prerna Singh	2022-07-27	13:11:06	20:58:35	07:47:29	1	2022-07-30	Prerna Singh	11	00:00:38	00:04:20
4.11	9										

[cloudera@quickstart ~]$ cat /tmp/Agent_data/agent_left_join.csv

1	Shivananda Sonwane	2022-07-30	15:35:29	17:39:39	02:04:10	69	2022-07-30	Shivananda Sonwane	4
00:01:14	00:16:53	5.0 1							
1	Shivananda Sonwane	2022-07-30	15:35:29	17:39:39	02:04:10	73	2022-07-29	Shivananda Sonwane	14
00:00:45	00:15:38	4.679							
1	Shivananda Sonwane	2022-07-30	15:35:29	17:39:39	02:04:10	214	2022-07-28	Shivananda Sonwane	5
00:00:31	00:38:04	5.0 4							
1	Shivananda Sonwane	2022-07-30	15:35:29	17:39:39	02:04:10	285	2022-07-27	Shivananda Sonwane	26
00:01:12	00:20:10	4.2218							
1	Shivananda Sonwane	2022-07-30	15:35:29	17:39:39	02:04:10	360	2022-07-26	Shivananda Sonwane	24
00:00:51	00:22:28	5.0 14							

[cloudera@quickstart ~]$ cat /tmp/Agent_data/agent_right_join.csv

16	Prerna Singh	2022-07-30	12:32:28	14:10:08	01:37:40	1	2022-07-30	Prerna Singh	11	00:00:38	00:04:20
4.11	9										
75	Prerna Singh	2022-07-29	17:47:06	21:03:44	03:16:37	1	2022-07-30	Prerna Singh	11	00:00:38	00:04:20
4.11	9										
91	Prerna Singh	2022-07-29	15:08:22	17:20:49	02:12:27	1	2022-07-30	Prerna Singh	11	00:00:38	00:04:20
4.11	9										
110	Prerna Singh	2022-07-29	12:08:23	12:11:35	00:03:11	1	2022-07-30	Prerna Singh	11	00:00:38	00:04:20
4.11	9										

336   Prerna Singh   2022-07-27   13:11:06   20:58:35   07:47:29   1   2022-07-30   Prerna Singh   11   00:00:38   00:04:20
4.11   9

Q17) Perform partitioning on top of the agent column and then on top of that perform bucketing for each partitioning.

Solution:

First set the below mentioned properties to be true.

hive> set hive.exec.dynamic.partition=true;

hive> set hive.exec.dynamic.patition.mode=nonstrict;

Create Partition_bucketed table:

hive> create table partition_bucketed_loging

   > (

   > s_no int,

   > date date,

   > login_time string,

   > logout_time string,

   > duration string

   >)

   > partitioned by (agent string)

   > clustered by(s_no)

   > into 4 buckets

   > row format delimited

   > fields terminated by ','

   > stored as textfile;

OK

Time taken: 0.156 seconds

Load data into Patition_bucketed table:

hive> insert overwrite table partition_bucketed_loging partition(agent) select s_no, date, login_time, logout_time, duration, agent from agent_loging;

Query ID = cloudera_20221102040808_6531ce58-de2e-4c4a-a582-331557eb8787

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_1667377030429_0010, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667377030429_0010/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667377030429_0010

Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 0

2022-11-02 04:08:54,135 Stage-1 map = 0%, reduce = 0%

2022-11-02 04:09:15,365 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 6.32 sec

MapReduce Total cumulative CPU time: 6 seconds 320 msec

Ended Job = job_1667377030429_0010

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/agent.db/partition_bucketed_loging/.hive-staging_hive_2022-11-02_04-08-38_814_1998316040267771961-1/-ext-10000

Loading data to table agent.partition_bucketed_loging partition (agent=null)

Time taken for load dynamic partitions : 11817

Loading partition {agent=Sudhanshu Kumar}

Loading partition {agent=Aditya Shinde}

Loading partition {agent=Suraj S Bilgi}

Loading partition {agent=Hrisikesh Neogi}

Loading partition {agent=Dibyanshu}

Loading partition {agent=Shiva Srivastava}

Loading partition {agent=Saikumarreddy N}

Loading partition {agent=Jaydeep Dixit}

Loading partition {agent=Ankitjha}

Loading partition {agent=Prabir Kumar Satapathy}

Loading partition {agent=Deepranjan Gupta}

Loading partition {agent=Khushboo Priya}

Loading partition {agent=Hyder Abbas}

Loading partition {agent=Amersh}

Loading partition {agent=Anurag Tiwari}

Loading partition {agent=Madhulika G}

Loading partition {agent=Aravind}

Loading partition {agent=Prateek _iot}

Loading partition {agent=Mithun S}

Loading partition {agent=Mahesh Sarade}

Loading partition {agent=Ameya Jain}

Loading partition {agent=Swati}

Loading partition {agent=Bharath}

Loading partition {agent=Sowmiya Sivakumar}

Loading partition {agent=Ineuron Intelligence}

Loading partition {agent=Prerna Singh}

Loading partition {agent=Wasim}

Loading partition {agent=Zeeshan}

Loading partition {agent=Manjunatha A}

Loading partition {agent=Nishtha Jain}

Loading partition {agent=Harikrishnan Shaji}

Loading partition {agent=Shubham Sharma}

Loading partition {agent=Boktiar Ahmed Bappy}

Loading partition {agent=Saurabh Shukla}

Loading partition {agent=Rishav Dash}

Loading partition {agent=Jawala Prakash}

Loading partition {agent=Mukesh}

Loading partition {agent=Tarun}

Loading partition {agent=Sanjeev Kumar}

Loading partition {agent=Maitry}

Loading partition {agent=Nandani Gupta}

Loading partition {agent=Aditya_iot}

Loading partition {agent=Chaitra K Hiremath}

Loading partition {agent=Shivan K}

Loading partition {agent=Shivananda Sonwane}

Loading partition {agent=Nitin M}

Loading partition {agent=Ishawant Kumar}

Loading partition {agent=Muskan Garg}

Loading partition {agent=Ayushi Mishra}

Time taken for adding to write entity : 31

Partition agent.partition_bucketed_loging{agent=Aditya Shinde} stats: [numFiles=1, numRows=1, totalSize=42, rawDataSize=41]

Partition agent.partition_bucketed_loging{agent=Aditya_iot} stats: [numFiles=1, numRows=9, totalSize=377, rawDataSize=368]

Partition agent.partition_bucketed_loging{agent=Amersh} stats: [numFiles=1, numRows=4, totalSize=168, rawDataSize=164]

Partition agent.partition_bucketed_loging{agent=Ameya Jain} stats: [numFiles=1, numRows=10, totalSize=420, rawDataSize=410]

Partition agent.partition_bucketed_loging{agent=Ankitjha} stats: [numFiles=1, numRows=4, totalSize=168, rawDataSize=164]

Partition agent.partition_bucketed_loging{agent=Anurag Tiwari} stats: [numFiles=1, numRows=37, totalSize=1553, rawDataSize=1516]

Partition agent.partition_bucketed_loging{agent=Aravind} stats: [numFiles=1, numRows=10, totalSize=420, rawDataSize=410]

Partition agent.partition_bucketed_loging{agent=Ayushi Mishra} stats: [numFiles=1, numRows=18, totalSize=755, rawDataSize=737]

Partition agent.partition_bucketed_loging{agent=Bharath} stats: [numFiles=1, numRows=9, totalSize=378, rawDataSize=369]

Partition agent.partition_bucketed_loging{agent=Boktiar Ahmed Bappy} stats: [numFiles=1, numRows=17, totalSize=709, rawDataSize=692]

Partition agent.partition_bucketed_loging{agent=Chaitra K Hiremath} stats: [numFiles=1, numRows=13, totalSize=543, rawDataSize=530]

Partition agent.partition_bucketed_loging{agent=Deepranjan Gupta} stats: [numFiles=1, numRows=58, totalSize=2433, rawDataSize=2375]

Partition agent.partition_bucketed_loging{agent=Dibyanshu} stats: [numFiles=1, numRows=208, totalSize=8719, rawDataSize=8511]

Partition agent.partition_bucketed_loging{agent=Harikrishnan Shaji} stats: [numFiles=1, numRows=23, totalSize=963, rawDataSize=940]

Partition agent.partition_bucketed_loging{agent=Hrisikesh Neogi} stats: [numFiles=1, numRows=37, totalSize=1544, rawDataSize=1507]

Partition agent.partition_bucketed_loging{agent=Hyder Abbas} stats: [numFiles=1, numRows=2, totalSize=84, rawDataSize=82]

Partition agent.partition_bucketed_loging{agent=Ineuron Intelligence} stats: [numFiles=1, numRows=1, totalSize=42, rawDataSize=41]

Partition agent.partition_bucketed_loging{agent=Ishawant Kumar} stats: [numFiles=1, numRows=49, totalSize=2052, rawDataSize=2003]

Partition agent.partition_bucketed_loging{agent=Jawala Prakash} stats: [numFiles=1, numRows=16, totalSize=668, rawDataSize=652]

Partition agent.partition_bucketed_loging{agent=Jaydeep Dixit} stats: [numFiles=1, numRows=11, totalSize=459, rawDataSize=448]

Partition agent.partition_bucketed_loging{agent=Khushboo Priya} stats: [numFiles=1, numRows=18, totalSize=752, rawDataSize=734]

Partition agent.partition_bucketed_loging{agent=Madhulika G} stats: [numFiles=1, numRows=17, totalSize=713, rawDataSize=696]

Partition agent.partition_bucketed_loging{agent=Mahesh Sarade} stats: [numFiles=1, numRows=36, totalSize=1509, rawDataSize=1473]

Partition agent.partition_bucketed_loging{agent=Maitry} stats: [numFiles=1, numRows=5, totalSize=210, rawDataSize=205]

Partition agent.partition_bucketed_loging{agent=Manjunatha A} stats: [numFiles=1, numRows=8, totalSize=333, rawDataSize=325]

Partition agent.partition_bucketed_loging{agent=Mithun S} stats: [numFiles=1, numRows=14, totalSize=586, rawDataSize=572]

Partition agent.partition_bucketed_loging{agent=Mukesh} stats: [numFiles=1, numRows=3, totalSize=124, rawDataSize=121]

Partition agent.partition_bucketed_loging{agent=Muskan Garg} stats: [numFiles=1, numRows=12, totalSize=503, rawDataSize=491]

Partition agent.partition_bucketed_loging{agent=Nandani Gupta} stats: [numFiles=1, numRows=11, totalSize=458, rawDataSize=447]

Partition agent.partition_bucketed_loging{agent=Nishtha Jain} stats: [numFiles=1, numRows=18, totalSize=754, rawDataSize=736]

Partition agent.partition_bucketed_loging{agent=Nitin M} stats: [numFiles=1, numRows=1, totalSize=42, rawDataSize=41]

Partition agent.partition_bucketed_loging{agent=Prabir Kumar Satapathy} stats: [numFiles=1, numRows=26, totalSize=1091, rawDataSize=1065]

Partition agent.partition_bucketed_loging{agent=Prateek _iot} stats: [numFiles=1, numRows=17, totalSize=711, rawDataSize=694]

Partition agent.partition_bucketed_loging{agent=Prerna Singh} stats: [numFiles=1, numRows=18, totalSize=753, rawDataSize=735]

Partition agent.partition_bucketed_loging{agent=Rishav Dash} stats: [numFiles=1, numRows=12, totalSize=504, rawDataSize=492]

Partition agent.partition_bucketed_loging{agent=Saikumarreddy N} stats: [numFiles=1, numRows=10, totalSize=420, rawDataSize=410]

Partition agent.partition_bucketed_loging{agent=Sanjeev Kumar} stats: [numFiles=1, numRows=20, totalSize=839, rawDataSize=819]

Partition agent.partition_bucketed_loging{agent=Saurabh Shukla} stats: [numFiles=1, numRows=40, totalSize=1680, rawDataSize=1640]

Partition agent.partition_bucketed_loging{agent=Shiva Srivastava} stats: [numFiles=1, numRows=15, totalSize=629, rawDataSize=614]

Partition agent.partition_bucketed_loging{agent=Shivan K} stats: [numFiles=1, numRows=36, totalSize=1506, rawDataSize=1470]

Partition agent.partition_bucketed_loging{agent=Shivananda Sonwane} stats: [numFiles=1, numRows=15, totalSize=625, rawDataSize=610]

Partition agent.partition_bucketed_loging{agent=Shubham Sharma} stats: [numFiles=1, numRows=35, totalSize=1469, rawDataSize=1434]

Partition agent.partition_bucketed_loging{agent=Sowmiya Sivakumar} stats: [numFiles=1, numRows=24, totalSize=1005, rawDataSize=981]

Partition agent.partition_bucketed_loging{agent=Sudhanshu Kumar} stats: [numFiles=1, numRows=11, totalSize=462, rawDataSize=451]

Partition agent.partition_bucketed_loging{agent=Suraj S Bilgi} stats: [numFiles=1, numRows=5, totalSize=206, rawDataSize=201]

Partition agent.partition_bucketed_loging{agent=Swati} stats: [numFiles=1, numRows=5, totalSize=210, rawDataSize=205]

Partition agent.partition_bucketed_loging{agent=Tarun} stats: [numFiles=1, numRows=1, totalSize=43, rawDataSize=42]

Partition agent.partition_bucketed_loging{agent=Wasim} stats: [numFiles=1, numRows=20, totalSize=840, rawDataSize=820]

Partition agent.partition_bucketed_loging{agent=Zeeshan} stats: [numFiles=1, numRows=10, totalSize=419, rawDataSize=409]

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1   Cumulative CPU: 6.32 sec   HDFS Read: 61313 HDFS Write: 45367 SUCCESS

Total MapReduce CPU Time Spent: 6 seconds 320 msec

OK

Time taken: 57.42 seconds