

```
select count(*) from collision where unix_timestamp('00:00','HH:mm')<=unix_timestamp(crashtime,'HH:mm') and
unix_timestamp('08:00','HH:mm')>=unix_timestamp(crashtime,'HH:mm');
```

```
1 row selected (21.677 seconds)
0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select count(*) from collision where unix_timestamp('00:00','HH:mm')<=unix_timestamp(crashtime,'HH:mm') and unix_timestamp('08:00','HH:mm')>=unix_timestamp(crashtime,'HH:mm')
INFO : Compiling command(queryId=hive_20220420155841_393a08c8-d532-4d47-9d9c-d562e134b41e); select count(*) from collision where unix_timestamp('00:00','HH:mm')<=unix_timestamp(crashtime,'HH:mm') and unix_timestamp('08:00','HH:mm')>=unix_timestamp(crashtime,'HH:mm')
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name: _c0, type:bigint, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220420155841_393a08c8-d532-4d47-9d9c-d562e134b41e); Time taken: 0.093 seconds
INFO : Executing command(queryId=hive_20220420155841_393a08c8-d532-4d47-9d9c-d562e134b41e); select count(*) from collision where unix_timestamp('00:00','HH:mm')<=unix_timestamp(crashtime,'HH:mm') and unix_timestamp('08:00','HH:mm')>=unix_timestamp(crashtime,'HH:mm')
WARN :
INFO : Query ID = hive_20220420155841_393a08c8-d532-4d47-9d9c-d562e134b41e
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:3
INFO : Submitting tokens for job: job_1648648882306_26942
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_1648648882306_26942/
INFO : Starting Job = job_1648648882306_26942, Tracking URL = http://horton.hpc.nyu.edu:8088/proxy/application_1648648882306_26942/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/hadoop job -kill job_1648648882306_26942
INFO : Hadoop job information for Stage-1: number of mappers: 3; number of reducers: 1
INFO : 2022-04-20 15:58:46.969 Stage-1 map = 0%, reduce = 0%
INFO : 2022-04-20 15:58:53.044 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 6.49 sec
INFO : 2022-04-20 15:58:54.061 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 16.15 sec
INFO : 2022-04-20 15:58:55.074 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 26.16 sec
INFO : 2022-04-20 15:59:01.141 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 30.07 sec
INFO : MapReduce Total cumulative CPU time: 30 seconds 70 msec
INFO : Ended Job = job_1648648882306_26942
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 3 Reduce: 1 Cumulative CPU: 30.07 sec HDFS Read: 404029308 HDFS Write: 106 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 30 seconds 70 msec
INFO : Completed executing command(queryId=hive_20220420155841_393a08c8-d532-4d47-9d9c-d562e134b41e); Time taken: 21.06 seconds
INFO : OK
+-----+
| _c0 |
+-----+
| 298753 |
+-----+
```

```
select count(*) from collision where unix_timestamp('08:00','HH:mm')<=unix_timestamp(crashtime,'HH:mm') and
unix_timestamp('16:00','HH:mm')>=unix_timestamp(crashtime,'HH:mm');
```

```
+-----+
| _c0 |
+-----+
| 884454 |
+-----+
```

```
0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select count(*) from collision where unix_timestamp('08:00','HH:mm')<=unix_timestamp(crashtime,'HH:mm') and unix_timestamp('16:00','HH:mm')>=unix_timestamp(crashtime,'HH:mm')
INFO : Compiling command(queryId=hive_20220420155931_4cd57891-8516-439c-a4dd-f0ca76287c9d); select count(*) from collision where unix_timestamp('08:00','HH:mm')<=unix_timestamp(crashtime,'HH:mm') and unix_timestamp('16:00','HH:mm')>=unix_timestamp(crashtime,'HH:mm')
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name: _c0, type:bigint, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220420155931_4cd57891-8516-439c-a4dd-f0ca76287c9d); Time taken: 0.1 seconds
INFO : Executing command(queryId=hive_20220420155931_4cd57891-8516-439c-a4dd-f0ca76287c9d); select count(*) from collision where unix_timestamp('08:00','HH:mm')<=unix_timestamp(crashtime,'HH:mm') and unix_timestamp('16:00','HH:mm')>=unix_timestamp(crashtime,'HH:mm')
WARN :
INFO : Query ID = hive_20220420155931_4cd57891-8516-439c-a4dd-f0ca76287c9d
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:3
INFO : Submitting tokens for job: job_1648648882306_26945
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_1648648882306_26945/
INFO : Starting Job = job_1648648882306_26945, Tracking URL = http://horton.hpc.nyu.edu:8088/proxy/application_1648648882306_26945/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/hadoop job -kill job_1648648882306_26945
INFO : Hadoop job information for Stage-1: number of mappers: 3; number of reducers: 1
INFO : 2022-04-20 15:59:37.415 Stage-1 map = 0%, reduce = 0%
INFO : 2022-04-20 15:59:42.482 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 6.33 sec
INFO : 2022-04-20 15:59:44.507 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 15.08 sec
INFO : 2022-04-20 15:59:45.518 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 24.84 sec
INFO : 2022-04-20 15:59:51.585 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 28.05 sec
INFO : MapReduce Total cumulative CPU time: 28 seconds 50 msec
INFO : Ended Job = job_1648648882306_26945
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 3 Reduce: 1 Cumulative CPU: 28.05 sec HDFS Read: 404029578 HDFS Write: 106 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 28 seconds 50 msec
INFO : Completed executing command(queryId=hive_20220420155931_4cd57891-8516-439c-a4dd-f0ca76287c9d); Time taken: 21.261 seconds
INFO : OK
+-----+
| _c0 |
+-----+
| 884454 |
+-----+
```

```
select count(*) from collision where unix_timestamp('16:00','HH:mm')<=unix_timestamp(crashTime,'HH:mm') and
unix_timestamp('24:00','HH:mm')>=unix_timestamp(crashTime,'HH:mm');
```

```
± 100 SELECTED (22.393 seconds)
|0: jdbc:hive2://hm-1.hpc.nyu.edu:10000/> select count(*) from collision where unix_timestamp('16:00','HH:mm')<=unix_timestamp(crashTime,'HH:mm') and unix_timestamp('24:00','HH:mm')>=unix_timestamp(crashTime,'HH:mm');
INFO : Compiling command(queryId=hive_20220420160022_8f3e582d-9bb9-4861-8498-352162473b37): select count(*) from collision where unix_timestamp('16:00','HH:mm')<=unix_timestamp(crashTime,'HH:mm') and unix_timestamp('24:00','HH:mm')>=unix_timestamp(crashTime,'HH:mm')
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldsSchemas:[FieldSchema(name:_c0, type:bigint, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220420160022_8f3e582d-9bb9-4861-8498-352162473b37); Time taken: 0.117 seconds
INFO : Executing command(queryId=hive_20220420160022_8f3e582d-9bb9-4861-8498-352162473b37): select count(*) from collision where unix_timestamp('16:00','HH:mm')<=unix_timestamp(crashTime,'HH:mm') and unix_timestamp('24:00','HH:mm')>=unix_timestamp(crashTime,'HH:mm')
WARN :
INFO : Query ID = hive_20220420160022_8f3e582d-9bb9-4861-8498-352162473b37
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:3
INFO : Submitting tokens for job: job_1648648882306_26946
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_1648648882306_26946/
INFO : Starting Job = job_1648648882306_26946, Tracking URL = http://horton.hpc.nyu.edu:8088/proxy/application_1648648882306_26946/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/hadoop job -kill job_1648648882306_26946
INFO : Hadoop job information for Stage-1: number of mappers: 3; number of reducers: 1
INFO : 2022-04-20 16:00:28,811 Stage-1 map = 0%, reduce = 0%
INFO : 2022-04-20 16:00:33,881 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 6.29 sec
INFO : 2022-04-20 16:00:35,911 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 15.63 sec
INFO : 2022-04-20 16:00:36,924 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 25.7 sec
INFO : 2022-04-20 16:00:41,986 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 30.8 sec
INFO : MapReduce Total cumulative CPU time: 30 seconds 800 msec
INFO : Ended Job = job_1648648882306_26946
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 3 Reduce: 1 Cumulative CPU: 30.8 sec HDFS Read: 404029308 HDFS Write: 106 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 30 seconds 800 msec
INFO : Completed executing command(queryId=hive_20220420160022_8f3e582d-9bb9-4861-8498-352162473b37); Time taken: 20.261 seconds
INFO : OK
+-----+
| _c0 |
+-----+
| 744209 |
+-----+
1 row selected (20.393 seconds)
```

Mta profile:

Create Table

```
create table mta(time STRING, subway INTEGER, subway_percent DOUBLE, ride INTEGER, ride_percent DOUBLE, LIRR INTEGER, LIRR_percent DOUBLE, Metro_North INTEGER, Metro_percent DOUBLE, Access INTEGER, Access_percent Float, traffic INTEGER, traffic_percent Double ) ROW FORMAT DELIMITED FIELDS TERMINATED BY ',' STORED AS TEXTFILE;
```

```
- -----
0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> create table mta(time STRING, subway IN
TEGER, subway_percent DOUBLE, ride INTEGER, ride_percent DOUBLE, LIRR INTEGER, L
IRR_percent DOUBLE, Metro_North INTEGER, Metro_percent DOUBLE, Access INTEGER, A
ccess_percent Float, traffic INTEGER, traffic_percent Double ) ROW FORMAT DELIMI
TED FIELDS TERMINATED BY ',' STORED AS TEXTFILE;
INFO  : Compiling command(queryId=hive_20220426160058_77b9b3af-93e8-400f-a463-db
7fb5229f9b): create table mta(time STRING, subway INTEGER, subway_percent DOUBLE
, ride INTEGER, ride_percent DOUBLE, LIRR INTEGER, LIRR_percent DOUBLE, Metro_No
rth INTEGER, Metro_percent DOUBLE, Access INTEGER, Access_percent Float, traffic
 INTEGER, traffic_percent Double ) ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
STORED AS TEXTFILE
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:null, properties:null)
INFO  : Completed compiling command(queryId=hive_20220426160058_77b9b3af-93e8-40
0f-a463-db7fb5229f9b); Time taken: 0.042 seconds
INFO  : Executing command(queryId=hive_20220426160058_77b9b3af-93e8-400f-a463-db
7fb5229f9b): create table mta(time STRING, subway INTEGER, subway_percent DOUBLE
, ride INTEGER, ride_percent DOUBLE, LIRR INTEGER, LIRR_percent DOUBLE, Metro_No
rth INTEGER, Metro_percent DOUBLE, Access INTEGER, Access_percent Float, traffic
 INTEGER, traffic_percent Double ) ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
STORED AS TEXTFILE
INFO  : Starting task [Stage-0:DDL] in serial mode
INFO  : Completed executing command(queryId=hive_20220426160058_77b9b3af-93e8-40
0f-a463-db7fb5229f9b); Time taken: 0.374 seconds
INFO  : OK
No rows affected (0.422 seconds)
0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> describe mta;
INFO  : Compiling command(queryId=hive_20220426160102_bcdb5366-9dfe-400e-be9f-78
3bb585da50): describe mta
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:col_name, t
ype:string, comment:from deserializer), FieldSchema(name:data_type, type:string,
comment:from deserializer), FieldSchema(name:comment, type:string, comment:from
deserializer)], properties:null)
INFO  : Completed compiling command(queryId=hive_20220426160102_bcdb5366-9dfe-40
0e-be9f-783bb585da50); Time taken: 0.047 seconds
INFO  : Executing command(queryId=hive_20220426160102_bcdb5366-9dfe-400e-be9f-78
3bb585da50): describe mta
INFO  : Starting task [Stage-0:DDL] in serial mode
INFO  : Completed executing command(queryId=hive_20220426160102_bcdb5366-9dfe-40
0e-be9f-783bb585da50); Time taken: 0.003 seconds
INFO  : OK
```

```

INFO  . completed executing command(queryId=hive_20220426160102_bcdb5366-9dfe-400e-be9f-783bb585da50); Time taken: 0.374 seconds
INFO  : OK
No rows affected (0.422 seconds)
[0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> describe mta;
INFO  : Compiling command(queryId=hive_20220426160102_bcdb5366-9dfe-400e-be9f-783bb585da50): describe mta
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:col_name, type:string, comment:from deserializer), FieldSchema(name:data_type, type:string, comment:from deserializer), FieldSchema(name:comment, type:string, comment:from deserializer)], properties:null)
INFO  : Completed compiling command(queryId=hive_20220426160102_bcdb5366-9dfe-400e-be9f-783bb585da50); Time taken: 0.047 seconds
INFO  : Executing command(queryId=hive_20220426160102_bcdb5366-9dfe-400e-be9f-783bb585da50): describe mta
INFO  : Starting task [Stage-0:DDL] in serial mode
INFO  : Completed executing command(queryId=hive_20220426160102_bcdb5366-9dfe-400e-be9f-783bb585da50); Time taken: 0.003 seconds
INFO  : OK
+-----+-----+-----+
| col_name | data_type | comment |
+-----+-----+-----+
| time     | string    |          |
| subway   | int       |          |
| subway_percent | double   |          |
| ride     | int       |          |
| ride_percent | double   |          |
| lIRR      | int       |          |
| lIRR_percent | double   |          |
| metro_north | int       |          |
| metro_percent | double   |          |
| access    | int       |          |
| access_percent | float   |          |
| traffic   | int       |          |
| traffic_percent | double   |          |
+-----+-----+-----+

```

Load data:

```
load data inpath '/user/th2160/hw9/daily.csv' into table mta;
```

```
line 11 character 1: not supported here (state=42000, code=40000,
0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> load data inpath '/user/th2160/hw9/daily.csv' into table mta;
INFO : Compiling command(queryId=hive_20220426160135_b9b82846-aeb6-4b1c-aa0e-5ab61f29d665): load data inpath '/user/th2160/hw9/daily.csv' into table mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:null, properties:null)
INFO : Completed compiling command(queryId=hive_20220426160135_b9b82846-aeb6-4b1c-aa0e-5ab61f29d665); Time taken: 0.05 seconds
INFO : Executing command(queryId=hive_20220426160135_b9b82846-aeb6-4b1c-aa0e-5ab61f29d665): load data inpath '/user/th2160/hw9/daily.csv' into table mta
INFO : Starting task [Stage-0:MOVE] in serial mode
INFO : Loading data to table th2160.mta from hdfs://horton.hpc.nyu.edu:8020/user/th2160/hw9/daily.csv
INFO : Starting task [Stage-1:STATS] in serial mode
INFO : Completed executing command(queryId=hive_20220426160135_b9b82846-aeb6-4b1c-aa0e-5ab61f29d665); Time taken: 0.365 seconds
INFO : OK
No rows affected (0.423 seconds)
```

Count lines to see whether data is correct:

```
select count(*) from mta;
```

```

INFO 1 rows directed (0.420 seconds)
[0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select count(*) from mta;      ]
INFO : Compiling command(queryId=hive_20220426160146_dbe58563-80bd-4237-a23e-8c
f60f135a67): select count(*) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:b
igint, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426160146_dbe58563-80bd-42
37-a23e-8cf60f135a67); Time taken: 0.219 seconds
INFO : Executing command(queryId=hive_20220426160146_dbe58563-80bd-4237-a23e-8c
f60f135a67): select count(*) from mta
WARN :
INFO : Query ID = hive_20220426160146_dbe58563-80bd-4237-a23e-8cf60f135a67
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33459
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/applicati
on_1648648882306_33459/
INFO : Starting Job = job_1648648882306_33459, Tracking URL = http://horton.hpc
.nyu.edu:8088/proxy/application_1648648882306_33459/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib
/hadoop/bin/hadoop job -kill job_1648648882306_33459
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of redu
cers: 1
INFO : 2022-04-26 16:01:51,572 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 16:01:57,656 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU
 1.85 sec
INFO : 2022-04-26 16:02:02,719 Stage-1 map = 100%,  reduce = 100%, Cumulative C
PU 6.92 sec
INFO : MapReduce Total cumulative CPU time: 6 seconds 920 msec
INFO : Ended Job = job_1648648882306_33459
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.92 sec  HDFS Read:
 75565 HDFS Write: 103 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 920 msec
INFO : Completed executing command(queryId=hive_20220426160146_dbe58563-80bd-42
37-a23e-8cf60f135a67); Time taken: 17.249 seconds
INFO : OK
+---+
| _c0 |
+---+
| 776 |
+---+
1 row selected (17.483 seconds)

```

Count missing value for each column :

```
select count(*) from mta where lirr is NULL;
```

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select count(*) from mta where lirr is ]
NULL;
INFO  : Compiling command(queryId=hive_20220426160244_fba24dba-2a82-4e95-9e40-30
a751272082): select count(*) from mta where lirr is NULL
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:b
igint, comment:null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20220426160244_fba24dba-2a82-4e
95-9e40-30a751272082); Time taken: 0.11 seconds
INFO  : Executing command(queryId=hive_20220426160244_fba24dba-2a82-4e95-9e40-30
a751272082): select count(*) from mta where lirr is NULL
WARN  :
INFO  : Query ID = hive_20220426160244_fba24dba-2a82-4e95-9e40-30a751272082
INFO  : Total jobs = 1
INFO  : Launching Job 1 out of 1
INFO  : Starting task [Stage-1:MAPRED] in serial mode
INFO  : Number of reduce tasks determined at compile time: 1
INFO  : In order to change the average load for a reducer (in bytes):
INFO  :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO  : In order to limit the maximum number of reducers:
INFO  :   set hive.exec.reducers.max=<number>
INFO  : In order to set a constant number of reducers:
INFO  :   set mapreduce.job.reduces=<number>
INFO  : number of splits:1
INFO  : Submitting tokens for job: job_1648648882306_33460
INFO  : Executing with tokens: []
INFO  : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/applicati
on_1648648882306_33460/
INFO  : Starting Job = job_1648648882306_33460, Tracking URL = http://horton.hpc
.ny.edu:8088/proxy/application_1648648882306_33460/
INFO  : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib
/hadoop/bin/hadoop job -kill job_1648648882306_33460
INFO  : Hadoop job information for Stage-1: number of mappers: 1; number of redu
cers: 1
INFO  : 2022-04-26 16:04:05,596 Stage-1 map = 0%,  reduce = 0%
INFO  : 2022-04-26 16:04:10,758 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU
 6.14 sec
INFO  : 2022-04-26 16:04:16,826 Stage-1 map = 100%,  reduce = 100%, Cumulative C
PU 10.27 sec
INFO  : MapReduce Total cumulative CPU time: 10 seconds 270 msec
INFO  : Ended Job = job_1648648882306_33460
INFO  : MapReduce Jobs Launched:
INFO  : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 10.27 sec    HDFS Read
: 76598 HDFS Write: 102 HDFS EC Read: 0 SUCCESS
INFO  : Total MapReduce CPU Time Spent: 10 seconds 270 msec
INFO  : Completed executing command(queryId=hive_20220426160244_fba24dba-2a82-4e
95-9e40-30a751272082); Time taken: 114.688 seconds
INFO  : OK
+---+
| _c0 |
+---+
| 33  |
+---+
1 row selected (114.809 seconds)

```

```
select count(*) from mta where lirr_percent is NULL;
```

```
0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select count(*) from mta where lirr_percent is NULL;
INFO : Compiling command(queryId=hive_20220426160755_35c98a64-91e8-48dd-8686-125937eba93e): select count(*) from mta where lirr_percent is NULL
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:b
igint, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426160755_35c98a64-91e8-48
dd-8686-125937eba93e); Time taken: 0.103 seconds
INFO : Executing command(queryId=hive_20220426160755_35c98a64-91e8-48dd-8686-12
5937eba93e): select count(*) from mta where lirr_percent is NULL
WARN :
INFO : Query ID = hive_20220426160755_35c98a64-91e8-48dd-8686-125937eba93e
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33467
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/applicati
on_1648648882306_33467/
INFO : Starting Job = job_1648648882306_33467, Tracking URL = http://horton.hpc
.ny.edu:8088/proxy/application_1648648882306_33467/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib
/hadoop/bin/hadoop job -kill job_1648648882306_33467
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of redu
cers: 1
INFO : 2022-04-26 16:08:02,083 Stage-1 map = 0%, reduce = 0%
INFO : 2022-04-26 16:08:08,165 Stage-1 map = 100%, reduce = 0%, Cumulative CPU
6.24 sec
INFO : 2022-04-26 16:08:14,235 Stage-1 map = 100%, reduce = 100%, Cumulative C
PU 9.07 sec
INFO : MapReduce Total cumulative CPU time: 9 seconds 70 msec
INFO : Ended Job = job_1648648882306_33467
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 9.07 sec HDFS Read:
76605 HDFS Write: 102 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 9 seconds 70 msec
INFO : Completed executing command(queryId=hive_20220426160755_35c98a64-91e8-48
dd-8686-125937eba93e); Time taken: 20.23 seconds
INFO : OK
+---+
| _c0 |
+---+
| 33 |
```

```

select count(*) from mta where time is NULL;

,code=40000,
0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select count(*) from mta where time is NULL;
INFO  : Compiling command(queryId=hive_20220426160848_071a3cc4-6610-4e4e-9e66-47cbf10a7541): select count(*) from mta where time is NULL
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:bigint, comment:null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20220426160848_071a3cc4-6610-4e4e-9e66-47cbf10a7541); Time taken: 0.085 seconds
INFO  : Executing command(queryId=hive_20220426160848_071a3cc4-6610-4e4e-9e66-47cbf10a7541): select count(*) from mta where time is NULL
WARN  :
INFO  : Query ID = hive_20220426160848_071a3cc4-6610-4e4e-9e66-47cbf10a7541
INFO  : Total jobs = 1
INFO  : Launching Job 1 out of 1
INFO  : Starting task [Stage-1:MAPRED] in serial mode
INFO  : Number of reduce tasks determined at compile time: 1
INFO  : In order to change the average load for a reducer (in bytes):
INFO  :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO  : In order to limit the maximum number of reducers:
INFO  :   set hive.exec.reducers.max=<number>
INFO  : In order to set a constant number of reducers:
INFO  :   set mapreduce.job.reduces=<number>
INFO  : number of splits:1
INFO  : Submitting tokens for job: job_1648648882306_33471
INFO  : Executing with tokens: []
INFO  : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_1648648882306_33471/
INFO  : Starting Job = job_1648648882306_33471, Tracking URL = http://horton.hpc.nyu.edu:8088/proxy/application_1648648882306_33471/
INFO  : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/hadoop job -kill job_1648648882306_33471
INFO  : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO  : 2022-04-26 16:08:54,076 Stage-1 map = 0%,  reduce = 0%
INFO  : 2022-04-26 16:09:00,152 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 6.74 sec
INFO  : 2022-04-26 16:09:06,222 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 11.57 sec
INFO  : MapReduce Total cumulative CPU time: 11 seconds 570 msec
INFO  : Ended Job = job_1648648882306_33471
INFO  : MapReduce Jobs Launched:
INFO  : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 11.57 sec    HDFS Read: 76524 HDFS Write: 101 HDFS EC Read: 0 SUCCESS
INFO  : Total MapReduce CPU Time Spent: 11 seconds 570 msec
INFO  : Completed executing command(queryId=hive_20220426160848_071a3cc4-6610-4e4e-9e66-47cbf10a7541); Time taken: 18.328 seconds
INFO  : OK
+---+
| _c0 |
+---+
| 0   |

```

```

select count(*) from mta where subway is NULL;
0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select count(*) from mta where subway is NULL;
INFO : Compiling command(queryId=hive_20220426160942_9c134603-bb89-4765-a0dd-b0ed204b941b): select count(*) from mta where subway is NULL
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:bigint, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426160942_9c134603-bb89-4765-a0dd-b0ed204b941b); Time taken: 0.086 seconds
INFO : Executing command(queryId=hive_20220426160942_9c134603-bb89-4765-a0dd-b0ed204b941b): select count(*) from mta where subway is NULL
WARN :
INFO : Query ID = hive_20220426160942_9c134603-bb89-4765-a0dd-b0ed204b941b
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33475
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_1648648882306_33475/
INFO : Starting Job = job_1648648882306_33475, Tracking URL = http://horton.hpc.nyu.edu:8088/proxy/application_1648648882306_33475/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/hadoop job -kill job_1648648882306_33475
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 16:09:48,098 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 16:09:53,164 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 6.43 sec
INFO : 2022-04-26 16:09:59,253 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 10.64 sec
INFO : MapReduce Total cumulative CPU time: 10 seconds 640 msec
INFO : Ended Job = job_1648648882306_33475
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 10.64 sec HDFS Read: 76598 HDFS Write: 101 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 10 seconds 640 msec
INFO : Completed executing command(queryId=hive_20220426160942_9c134603-bb89-4765-a0dd-b0ed204b941b); Time taken: 18.129 seconds
INFO : OK
+---+
| _c0 |
+---+
| 1   |
+---+

```

```
select count(*) from mta where subway_percent is NULL;
```

```
+--+
[0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select count(*) from mta where subway_p]
ercent is NULL;
INFO  : Compiling command(queryId=hive_20220426161038_48ab4282-78a7-408e-96eb-a0
d7e95e5088): select count(*) from mta where subway_percent is NULL
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:b
igint, comment:null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20220426161038_48ab4282-78a7-40
8e-96eb-a0d7e95e5088); Time taken: 0.09 seconds
INFO  : Executing command(queryId=hive_20220426161038_48ab4282-78a7-408e-96eb-a0
d7e95e5088): select count(*) from mta where subway_percent is NULL
WARN  :
INFO  : Query ID = hive_20220426161038_48ab4282-78a7-408e-96eb-a0d7e95e5088
INFO  : Total jobs = 1
INFO  : Launching Job 1 out of 1
INFO  : Starting task [Stage-1:MAPRED] in serial mode
INFO  : Number of reduce tasks determined at compile time: 1
INFO  : In order to change the average load for a reducer (in bytes):
INFO  :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO  : In order to limit the maximum number of reducers:
INFO  :   set hive.exec.reducers.max=<number>
INFO  : In order to set a constant number of reducers:
INFO  :   set mapreduce.job.reduces=<number>
INFO  : number of splits:1
INFO  : Submitting tokens for job: job_1648648882306_33476
INFO  : Executing with tokens: []
INFO  : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/applicati
on_1648648882306_33476/
INFO  : Starting Job = job_1648648882306_33476, Tracking URL = http://horton.hpc
.ny.edu:8088/proxy/application_1648648882306_33476/
INFO  : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib
/hadoop/bin/hadoop job -kill job_1648648882306_33476
INFO  : Hadoop job information for Stage-1: number of mappers: 1; number of redu
cers: 1
INFO  : 2022-04-26 16:10:44,372 Stage-1 map = 0%,  reduce = 0%
INFO  : 2022-04-26 16:10:49,440 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU
 5.61 sec
INFO  : 2022-04-26 16:10:55,510 Stage-1 map = 100%,  reduce = 100%, Cumulative C
PU 8.95 sec
INFO  : MapReduce Total cumulative CPU time: 8 seconds 950 msec
INFO  : Ended Job = job_1648648882306_33476
INFO  : MapReduce Jobs Launched:
INFO  : Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 8.95 sec  HDFS Read:
 76606  HDFS Write: 101  HDFS EC Read: 0  SUCCESS
INFO  : Total MapReduce CPU Time Spent: 8 seconds 950 msec
INFO  : Completed executing command(queryId=hive_20220426161038_48ab4282-78a7-40
8e-96eb-a0d7e95e5088); Time taken: 18.131 seconds
INFO  : OK
+----+
| _c0 |
+----+
| 1   |
+----+
```

```

select count(*) from mta where ride is NULL;
+-----+
| _c0 |
+-----+
| 1   |
+-----+
[0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select count(*) from mta where ride is NULL;
INFO : Compiling command(queryId=hive_20220426161246_d6b74e81-4372-44ed-9003-621b2d991ef6): select count(*) from mta where ride is NULL
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:bigint, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426161246_d6b74e81-4372-44ed-9003-621b2d991ef6); Time taken: 0.103 seconds
INFO : Executing command(queryId=hive_20220426161246_d6b74e81-4372-44ed-9003-621b2d991ef6): select count(*) from mta where ride is NULL
WARN :
INFO : Query ID = hive_20220426161246_d6b74e81-4372-44ed-9003-621b2d991ef6
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :     set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :     set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :     set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33483
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_1648648882306_33483/
INFO : Starting Job = job_1648648882306_33483, Tracking URL = http://horton.hpc.nyu.edu:8088/proxy/application_1648648882306_33483/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/hadoop job -kill job_1648648882306_33483
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 16:12:52,315 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 16:14:05,586 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 16:14:09,632 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 5.84 sec
INFO : 2022-04-26 16:14:15,702 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 8.73 sec
INFO : MapReduce Total cumulative CPU time: 8 seconds 730 msec
INFO : Ended Job = job_1648648882306_33483
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 8.73 sec  HDFS Read: 76516 HDFS Write: 101 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 8 seconds 730 msec
INFO : Completed executing command(queryId=hive_20220426161246_d6b74e81-4372-44ed-9003-621b2d991ef6); Time taken: 90.393 seconds
INFO : OK
+-----+
| _c0 |
+-----+
| 1   |
+-----+

```

```

select count(*) from mta where ride_percent is NULL;
[0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select count(*) from mta where ride_percent is NULL;
INFO : Compiling command(queryId=hive_20220426162215_3b000f76-0e95-42f0-b093-e8ec323aa823): select count(*) from mta where ride_percent is NULL
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:bigint, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426162215_3b000f76-0e95-42f0-b093-e8ec323aa823); Time taken: 0.124 seconds
INFO : Executing command(queryId=hive_20220426162215_3b000f76-0e95-42f0-b093-e8ec323aa823): select count(*) from mta where ride_percent is NULL
WARN :
INFO : Query ID = hive_20220426162215_3b000f76-0e95-42f0-b093-e8ec323aa823
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33504
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_1648648882306_33504/
INFO : Starting Job = job_1648648882306_33504, Tracking URL = http://horton.hpc.nyu.edu:8088/proxy/application_1648648882306_33504/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/hadoop job -kill job_1648648882306_33504
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 16:22:20,971 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 16:22:27,047 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 6.1 sec
INFO : 2022-04-26 16:22:33,120 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 9.09 sec
INFO : MapReduce Total cumulative CPU time: 9 seconds 90 msec
INFO : Ended Job = job_1648648882306_33504
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 9.09 sec HDFS Read: 76524 HDFS Write: 101 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 9 seconds 90 msec
INFO : Completed executing command(queryId=hive_20220426162215_3b000f76-0e95-42f0-b093-e8ec323aa823); Time taken: 18.353 seconds
INFO : OK
+---+
| _c0 |
+---+
| 1 |
+---+
1 row selected (18.488 seconds)

select count(*) from mta where metro_north is NULL;
```

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000/> select count(*) from mta where metro_no]
rth is NULL;
INFO : Compiling command(queryId=hive_20220426162316_17d08492-a2d1-4b07-95c1-a6
55febfbabe8): select count(*) from mta where metro_north is NULL
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:b
igint, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426162316_17d08492-a2d1-4b
07-95c1-a655febfbabe8); Time taken: 0.101 seconds
INFO : Executing command(queryId=hive_20220426162316_17d08492-a2d1-4b07-95c1-a6
55febfbabe8): select count(*) from mta where metro_north is NULL
WARN :
INFO : Query ID = hive_20220426162316_17d08492-a2d1-4b07-95c1-a655febfbabe8
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33508
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/applicati
on_1648648882306_33508/
INFO : Starting Job = job_1648648882306_33508, Tracking URL = http://horton.hpc
.ny.edu:8088/proxy/application_1648648882306_33508/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib
/hadoop/bin/hadoop job -kill job_1648648882306_33508
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of redu
cers: 1
INFO : 2022-04-26 16:23:21,675 Stage-1 map = 0%, reduce = 0%
INFO : 2022-04-26 16:23:27,754 Stage-1 map = 100%, reduce = 0%, Cumulative CPU
5.72 sec
INFO : 2022-04-26 16:23:33,823 Stage-1 map = 100%, reduce = 100%, Cumulative C
PU 10.13 sec
INFO : MapReduce Total cumulative CPU time: 10 seconds 130 msec
INFO : Ended Job = job_1648648882306_33508
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 10.13 sec HDFS Read
: 76590 HDFS Write: 102 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 10 seconds 130 msec
INFO : Completed executing command(queryId=hive_20220426162316_17d08492-a2d1-4b
07-95c1-a655febfbabe8); Time taken: 18.356 seconds
INFO : OK
+---+
| _c0 |
+---+
| 33 |
+---+
1 row selected (18.47 seconds)

```

```

select count(*) from mta where metro_percent is NULL;
0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select count(*) from mta where metro_percent is NULL;
INFO : Compiling command(queryId=hive_20220426162411_f821a66a-97e1-4c18-bced-ecd27ca6e070): select count(*) from mta where metro_percent is NULL
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:bigint, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426162411_f821a66a-97e1-4c18-bced-ecd27ca6e070); Time taken: 0.099 seconds
INFO : Executing command(queryId=hive_20220426162411_f821a66a-97e1-4c18-bced-ecd27ca6e070): select count(*) from mta where metro_percent is NULL
WARN :
INFO : Query ID = hive_20220426162411_f821a66a-97e1-4c18-bced-ecd27ca6e070
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33516
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_1648648882306_33516/
INFO : Starting Job = job_1648648882306_33516, Tracking URL = http://horton.hpc.nyu.edu:8088/proxy/application_1648648882306_33516/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/hadoop job -kill job_1648648882306_33516
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 16:24:17,218 Stage-1 map = 0%, reduce = 0%
INFO : 2022-04-26 16:24:24,319 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 4.68 sec
INFO : 2022-04-26 16:24:29,382 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 7.7 sec
INFO : MapReduce Total cumulative CPU time: 7 seconds 700 msec
INFO : Ended Job = job_1648648882306_33516
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 7.7 sec HDFS Read: 76522 HDFS Write: 102 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 7 seconds 700 msec
INFO : Completed executing command(queryId=hive_20220426162411_f821a66a-97e1-4c18-bced-ecd27ca6e070); Time taken: 19.355 seconds
INFO : OK
+---+
| _c0 |
+---+
| 33  |
+---+
1 row selected (19.466 seconds)

```

```

select count(*) from mta where access is NULL;
0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select count(*) from mta where access :s NULL;
INFO  : Compiling command(queryId=hive_20220426162653_4fcf554e-4de2-47f5-b0e9-61c65fad9c5c): select count(*) from mta where access is NULL
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:int, comment:null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20220426162653_4fcf554e-4de2-47f5-b0e9-6bc65fad9c5c); Time taken: 0.098 seconds
INFO  : Executing command(queryId=hive_20220426162653_4fcf554e-4de2-47f5-b0e9-61c65fad9c5c): select count(*) from mta where access is NULL
WARN  :
INFO  : Query ID = hive_20220426162653_4fcf554e-4de2-47f5-b0e9-6bc65fad9c5c
INFO  : Total jobs = 1
INFO  : Launching Job 1 out of 1
INFO  : Starting task [Stage-1:MAPRED] in serial mode
INFO  : Number of reduce tasks determined at compile time: 1
INFO  : In order to change the average load for a reducer (in bytes):
INFO  :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO  : In order to limit the maximum number of reducers:
INFO  :   set hive.exec.reducers.max=<number>
INFO  : In order to set a constant number of reducers:
INFO  :   set mapreduce.job.reduces=<number>
INFO  : number of splits:1
INFO  : Submitting tokens for job: job_1648648882306_33522
INFO  : Executing with tokens: []
INFO  : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_1648648882306_33522/
INFO  : Starting Job = job_1648648882306_33522, Tracking URL = http://horton.hpc.nyu.edu:8088/proxy/application_1648648882306_33522/
INFO  : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/hadoop job -kill job_1648648882306_33522
INFO  : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO  : 2022-04-26 16:29:01,966 Stage-1 map = 0%,  reduce = 0%
INFO  : 2022-04-26 16:29:11,085 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 5.78 sec
INFO  : 2022-04-26 16:29:16,152 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 8.58 sec
INFO  : MapReduce Total cumulative CPU time: 8 seconds 580 msec
INFO  : Ended Job = job_1648648882306_33522
INFO  : MapReduce Jobs Launched:
INFO  : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 8.58 sec    HDFS Read: 76589 HDFS Write: 101 HDFS EC Read: 0 SUCCESS
INFO  : Total MapReduce CPU Time Spent: 8 seconds 580 msec
INFO  : Completed executing command(queryId=hive_20220426162653_4fcf554e-4de2-47f5-b0e9-6bc65fad9c5c); Time taken: 143.273 seconds
INFO  : OK
+----+
| _c0 |
+----+
| 5   |
+----+

```

```

select count(*) from mta where access_percent is NULL;
[0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select count(*) from mta where access_p]
ercent is NULL;
INFO  : Compiling command(queryId=hive_20220426163008_c5582d36-8129-4cb2-a429-56
9bee92120b): select count(*) from mta where access_percent is NULL
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:b
igint, comment:null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20220426163008_c5582d36-8129-4c
b2-a429-569bee92120b); Time taken: 0.1 seconds
INFO  : Executing command(queryId=hive_20220426163008_c5582d36-8129-4cb2-a429-56
9bee92120b): select count(*) from mta where access_percent is NULL
WARN  :
INFO  : Query ID = hive_20220426163008_c5582d36-8129-4cb2-a429-569bee92120b
INFO  : Total jobs = 1
INFO  : Launching Job 1 out of 1
INFO  : Starting task [Stage-1:MAPRED] in serial mode
INFO  : Number of reduce tasks determined at compile time: 1
INFO  : In order to change the average load for a reducer (in bytes):
INFO  :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO  : In order to limit the maximum number of reducers:
INFO  :   set hive.exec.reducers.max=<number>
INFO  : In order to set a constant number of reducers:
INFO  :   set mapreduce.job.reduces=<number>
INFO  : number of splits:1
INFO  : Submitting tokens for job: job_1648648882306_33525
INFO  : Executing with tokens: []
INFO  : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/applicati
on_1648648882306_33525/
INFO  : Starting Job = job_1648648882306_33525, Tracking URL = http://horton.hpc
.ny.edu:8088/proxy/application_1648648882306_33525/
INFO  : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib
/hadoop/bin/hadoop job -kill job_1648648882306_33525
INFO  : Hadoop job information for Stage-1: number of mappers: 1; number of redu
cers: 1
INFO  : 2022-04-26 16:30:19,685 Stage-1 map = 0%,  reduce = 0%
INFO  : 2022-04-26 16:30:28,795 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU
 5.92 sec
INFO  : 2022-04-26 16:30:36,889 Stage-1 map = 100%,  reduce = 100%, Cumulative C
PU 8.5 sec
INFO  : MapReduce Total cumulative CPU time: 8 seconds 500 msec
INFO  : Ended Job = job_1648648882306_33525
INFO  : MapReduce Jobs Launched:
INFO  : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 8.5 sec    HDFS Read:
76428 HDFS Write: 101 HDFS EC Read: 0 SUCCESS
INFO  : Total MapReduce CPU Time Spent: 8 seconds 500 msec
INFO  : Completed executing command(queryId=hive_20220426163008_c5582d36-8129-4c
b2-a429-569bee92120b); Time taken: 30.337 seconds
INFO  : OK
+---+
| _c0 |
+---+
| 5   |
+---+
1 row selected (30.449 seconds)

```

```

select count(*) from mta where traffic is NULL;
[0: jdbc:hive2://hm-1.hpc.nyu.edu:10000] > select count(*) from mta where traffic :
is NULL;
INFO : Compiling command(queryId=hive_20220426163419_ecbc9027-aef2-49bf-b0b3-fc
0aa2ece59a): select count(*) from mta where traffic is NULL
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:b
igint, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426163419_ecbc9027-aef2-49
bf-b0b3-fc0aa2ece59a); Time taken: 0.116 seconds
INFO : Executing command(queryId=hive_20220426163419_ecbc9027-aef2-49bf-b0b3-fc
0aa2ece59a): select count(*) from mta where traffic is NULL
WARN :
INFO : Query ID = hive_20220426163419_ecbc9027-aef2-49bf-b0b3-fc0aa2ece59a
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33535
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/applicati
on_1648648882306_33535/
INFO : Starting Job = job_1648648882306_33535, Tracking URL = http://horton.hpc
.nyu.edu:8088/proxy/application_1648648882306_33535/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib
/hadoop/bin/hadoop job -kill job_1648648882306_33535
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of redu
cers: 1
INFO : 2022-04-26 16:34:31,825 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 16:34:36,893 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU
 5.53 sec
INFO : 2022-04-26 16:34:42,982 Stage-1 map = 100%,  reduce = 100%, Cumulative C
PU 9.93 sec
INFO : MapReduce Total cumulative CPU time: 9 seconds 930 msec
INFO : Ended Job = job_1648648882306_33535
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 9.93 sec    HDFS Read:
 76598 HDFS Write: 101 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 9 seconds 930 msec
INFO : Completed executing command(queryId=hive_20220426163419_ecbc9027-aef2-49
bf-b0b3-fc0aa2ece59a); Time taken: 26.434 seconds
INFO : OK
+---+
| _c0 |
+---+
| 1   |
+---+
1 row selected (26.561 seconds)

```

```

select count(*) from mta where traffic_percent is NULL;
[0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select count(*) from mta where traffic_
percent is NULL;
INFO : Compiling command(queryId=hive_20220426163535_334cdb3b-407d-466f-acbc-d5
34aa6ef21a): select count(*) from mta where traffic_percent is NULL
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:b
igint, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426163535_334cdb3b-407d-46
6f-acbc-d534aa6ef21a); Time taken: 0.118 seconds
INFO : Executing command(queryId=hive_20220426163535_334cdb3b-407d-466f-acbc-d5
34aa6ef21a): select count(*) from mta where traffic_percent is NULL
WARN :
INFO : Query ID = hive_20220426163535_334cdb3b-407d-466f-acbc-d534aa6ef21a
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33538
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/applicati
on_1648648882306_33538/
INFO : Starting Job = job_1648648882306_33538, Tracking URL = http://horton.hpc
.nyu.edu:8088/proxy/application_1648648882306_33538/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib
/hadoop/bin/hadoop job -kill job_1648648882306_33538
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of redu
cers: 1
INFO : 2022-04-26 16:36:02,148 Stage-1 map = 0%, reduce = 0%
INFO : 2022-04-26 16:36:27,436 Stage-1 map = 100%, reduce = 0%, Cumulative CPU
5.57 sec
INFO : 2022-04-26 16:36:37,555 Stage-1 map = 100%, reduce = 100%, Cumulative C
PU 8.46 sec
INFO : MapReduce Total cumulative CPU time: 8 seconds 460 msec
INFO : Ended Job = job_1648648882306_33538
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 8.46 sec HDFS Read:
76598 HDFS Write: 101 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 8 seconds 460 msec
INFO : Completed executing command(queryId=hive_20220426163535_334cdb3b-407d-46
6f-acbc-d534aa6ef21a); Time taken: 62.668 seconds
INFO : OK
+----+
| _c0 |
+----+
| 1 |
+----+
1 row selected (62.799 seconds)
a. idba:hive2://hm-1.hpc.nyu.edu:10000/>
```

```
insert overwrite into mta select * from mta where lirr is not NULL;
```

```
insert overwrite table mta select * from mta where subway is not NULL and subway_percent is not NULL and ride is not NULL and ride_percent is not NULL and lirr_percent is not NULL and metro_north is not NULL and metro_percent is not NULL and access is not NULL and access_percent is not NULL and traffic is not NULL and traffic_percent is not NULL;
```

```
0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> insert overwrite table mta select * from mta where subway is not NULL and subway_percent is not NULL and ride is not NULL and ride_percent is not NULL and lirr_percent is not NULL and metro_north is not NULL and metro_percent is not NULL and access is not NULL and access_percent is not NULL and traffic is not NULL and traffic_percent is not NULL;
INFO : Compiling command(queryId=hive_20220426170109_ecae6fb4-d120-42a0-985d-de13ee92cab8): insert overwrite table mta select * from mta where subway is not NULL and subway_percent is not NULL and ride is not NULL and ride_percent is not NULL and lirr_percent is not NULL and metro_no rth is not NULL and metro_percent is not NULL and access is not NULL and access_percent is not NULL and traffic is not NULL and traffic_percent is not NULL
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:mta.time, type:string, comment:null), FieldSchema(name:mta.subway, type:int, comment:null), FieldSchema(name:mta.subway_p ercent, type:double, comment:null), FieldSchema(name:mta.ride, type:int, comment:null), FieldSc hema(name:mta.ride_percent, type:double, comment:null), FieldSchema(name:mta.lirr, type:int, co mment:null), FieldSchema(name:mta.lirr_percent, type:double, comment:null), FieldSchema(name:mt a.metro_north, type:int, comment:null), FieldSchema(name:mta.metro_percent, type:double, commen t:null), FieldSchema(name:mta.access, type:int, comment:null), FieldSchema(name:mta.access_perc ent, type:float, comment:null), FieldSchema(name:mta.traffic, type:int, comment:null), FieldSch ema(name:mta.traffic_percent, type:double, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426170109_ecae6fb4-d120-42a0-985d-de13ee9 2cab8); Time taken: 0.091 seconds
INFO : Executing command(queryId=hive_20220426170109_ecae6fb4-d120-42a0-985d-de13ee92cab8): insert overwrite table mta select * from mta where subway is not NULL and subway_percent is not N ULL and ride is not NULL and ride_percent is not NULL and lirr_percent is not NULL and metro_no rth is not NULL and metro_percent is not NULL and access is not NULL and access_percent is not NULL and traffic is not NULL and traffic_percent is not NULL
WARN :
INFO : Query ID = hive_20220426170109_ecae6fb4-d120-42a0-985d-de13ee92cab8
INFO : Total jobs = 3
INFO : Launching Job 1 out of 3
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks is set to 0 since there's no reduce operator
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33568
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230 6_33568/
INFO : Starting Job = job_1648648882306_33568, Tracking URL = http://horton.hpc.nyu.edu:8088/p roxy/application_1648648882306_33568/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had oop job -kill job_1648648882306_33568
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 0
INFO : 2022-04-26 17:01:14,709 Stage-1 map = 0%,  reduce = 0%
```

```
t:null), FieldSchema(name:mta.access, type:int, comment:null), FieldSchema(name:mta.access_percent, type:float, comment:null), FieldSchema(name:mta.traffic, type:int, comment:null), FieldSchema(name:mta.traffic_percent, type:double, comment:null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20220426170109_ecae6fb4-d120-42a0-985d-de13ee92cab8); Time taken: 0.091 seconds
INFO  : Executing command(queryId=hive_20220426170109_ecae6fb4-d120-42a0-985d-de13ee92cab8): insert overwrite table mta select * from mta where subway is not NULL and subway_percent is not NULL and ride is not NULL and ride_percent is not NULL and lirr_percent is not NULL and metro_no_rth is not NULL and metro_percent is not NULL and access is not NULL and access_percent is not NULL and traffic is not NULL and traffic_percent is not NULL
WARN  :
INFO  : Query ID = hive_20220426170109_ecae6fb4-d120-42a0-985d-de13ee92cab8
INFO  : Total jobs = 3
INFO  : Launching Job 1 out of 3
INFO  : Starting task [Stage-1:MAPRED] in serial mode
INFO  : Number of reduce tasks is set to 0 since there's no reduce operator
INFO  : number of splits:1
INFO  : Submitting tokens for job: job_1648648882306_33568
INFO  : Executing with tokens: []
INFO  : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_1648648882306_33568/
INFO  : Starting Job = job_1648648882306_33568, Tracking URL = http://horton.hpc.nyu.edu:8088/proxy/application_1648648882306_33568/
INFO  : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/hadoop job -kill job_1648648882306_33568
INFO  : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 0
INFO  : 2022-04-26 17:01:14,709 Stage-1 map = 0%,  reduce = 0%
INFO  : 2022-04-26 17:01:20,786 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 3.37 sec
INFO  : MapReduce Total cumulative CPU time: 3 seconds 370 msec
INFO  : Ended Job = job_1648648882306_33568
INFO  : Starting task [Stage-7:CONDITIONAL] in serial mode
INFO  : Stage-4 is selected by condition resolver.
INFO  : Stage-3 is filtered out by condition resolver.
INFO  : Stage-5 is filtered out by condition resolver.
INFO  : Starting task [Stage-4:MOVE] in serial mode
INFO  : Moving data to directory hdfs://horton.hpc.nyu.edu:8020/user/hive/warehouse/th2160.db/mta/.hive-staging_hive_2022-04-26_17-01-09_545_5134606742165518503-2709/-ext-10000 from hdfs://horton.hpc.nyu.edu:8020/user/hive/warehouse/th2160.db/mta/.hive-staging_hive_2022-04-26_17-01-09_545_5134606742165518503-2709/-ext-10002
INFO  : Starting task [Stage-0:MOVE] in serial mode
INFO  : Loading data to table th2160.mta from hdfs://horton.hpc.nyu.edu:8020/user/hive/warehouse/th2160.db/mta/.hive-staging_hive_2022-04-26_17-01-09_545_5134606742165518503-2709/-ext-10000
INFO  : Starting task [Stage-2:STATS] in serial mode
INFO  : MapReduce Jobs Launched:
INFO  : Stage-Stage-1: Map: 1   Cumulative CPU: 3.37 sec   HDFS Read: 69422 HDFS Write: 62604 HDFS EC Read: 0 SUCCESS
INFO  : Total MapReduce CPU Time Spent: 3 seconds 370 msec
INFO  : Completed executing command(queryId=hive_20220426170109_ecae6fb4-d120-42a0-985d-de13ee92cab8); Time taken: 12.695 seconds
INFO  : OK
738 rows affected (12.794 seconds)
```

reflected

```
[0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select count(*) from mta;
INFO  : Compiling command(queryId=hive_20220426170135_1e532936-6b4e-4010-be77-ff
c2de02f6fe): select count(*) from mta
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:b
igint, comment:null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20220426170135_1e532936-6b4e-40
10-be77-ffc2de02f6fe); Time taken: 0.191 seconds
INFO  : Executing command(queryId=hive_20220426170135_1e532936-6b4e-4010-be77-ff
c2de02f6fe): select count(*) from mta
WARN  :
INFO  : Query ID = hive_20220426170135_1e532936-6b4e-4010-be77-ffc2de02f6fe
INFO  : Total jobs = 1
INFO  : Launching Job 1 out of 1
INFO  : Starting task [Stage-1:MAPRED] in serial mode
INFO  : Number of reduce tasks determined at compile time: 1
INFO  : In order to change the average load for a reducer (in bytes):
INFO  :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO  : In order to limit the maximum number of reducers:
INFO  :   set hive.exec.reducers.max=<number>
INFO  : In order to set a constant number of reducers:
INFO  :   set mapreduce.job.reduces=<number>
INFO  : number of splits:1
INFO  : Submitting tokens for job: job_1648648882306_33573
INFO  : Executing with tokens: []
INFO  : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/applicati
on_1648648882306_33573/
INFO  : Starting Job = job_1648648882306_33573, Tracking URL = http://horton.hpc
.ny.edu:8088/proxy/application_1648648882306_33573/
INFO  : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib
/hadoop/bin/hadoop job -kill job_1648648882306_33573
INFO  : Hadoop job information for Stage-1: number of mappers: 1; number of redu
cers: 1
INFO  : 2022-04-26 17:01:40,346 Stage-1 map = 0%,  reduce = 0%
INFO  : 2022-04-26 17:01:46,422 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU
 1.79 sec
INFO  : 2022-04-26 17:01:52,491 Stage-1 map = 100%,  reduce = 100%, Cumulative C
PU 6.25 sec
INFO  : MapReduce Total cumulative CPU time: 6 seconds 250 msec
INFO  : Ended Job = job_1648648882306_33573
INFO  : MapReduce Jobs Launched:
INFO  : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.25 sec    HDFS Read:
 72642 HDFS Write: 103 HDFS EC Read: 0 SUCCESS
INFO  : Total MapReduce CPU Time Spent: 6 seconds 250 msec
INFO  : Completed executing command(queryId=hive_20220426170135_1e532936-6b4e-40
10-be77-ffc2de02f6fe); Time taken: 18.263 seconds
INFO  : OK
+---+
| _c0 |
+---+
| 738 |
+---+
1 row selected (18.467 seconds)
```

```

select max(subway) from mta;
[0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select max(subway) from mta;          ]
INFO  : Compiling command(queryId=hive_20220426170636_a1c420ab-955f-4621-b4c0-d47b26102755): se
lect max(subway) from mta
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:int, comment:nul
l)], properties:null)
INFO  : Completed compiling command(queryId=hive_20220426170636_a1c420ab-955f-4621-b4c0-d47b261
02755); Time taken: 0.113 seconds
INFO  : Executing command(queryId=hive_20220426170636_a1c420ab-955f-4621-b4c0-d47b26102755): se
lect max(subway) from mta
WARN  :
INFO  : Query ID = hive_20220426170636_a1c420ab-955f-4621-b4c0-d47b26102755
INFO  : Total jobs = 1
INFO  : Launching Job 1 out of 1
INFO  : Starting task [Stage-1:MAPRED] in serial mode
INFO  : Number of reduce tasks determined at compile time: 1
INFO  : In order to change the average load for a reducer (in bytes):
INFO  :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO  : In order to limit the maximum number of reducers:
INFO  :   set hive.exec.reducers.max=<number>
INFO  : In order to set a constant number of reducers:
INFO  :   set mapreduce.job.reduces=<number>
INFO  : number of splits:1
INFO  : Submitting tokens for job: job_1648648882306_33589
INFO  : Executing with tokens: []
INFO  : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33589/
INFO  : Starting Job = job_1648648882306_33589, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33589/
INFO  : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33589
INFO  : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO  : 2022-04-26 17:06:43,557 Stage-1 map = 0%,  reduce = 0%
INFO  : 2022-04-26 17:06:48,633 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.84 sec
INFO  : 2022-04-26 17:06:53,689 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 4.8 sec
INFO  : MapReduce Total cumulative CPU time: 4 seconds 800 msec
INFO  : Ended Job = job_1648648882306_33589
INFO  : MapReduce Jobs Launched:
INFO  : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.8 sec    HDFS Read: 72818 HDFS Writ
e: 107 HDFS EC Read: 0 SUCCESS
INFO  : Total MapReduce CPU Time Spent: 4 seconds 800 msec
INFO  : Completed executing command(queryId=hive_20220426170636_a1c420ab-955f-4621-b4c0-d47b261
02755); Time taken: 17.708 seconds
INFO  : OK
+-----+
| _c0  |
+-----+
| 3440822 |
+-----+
1 row selected (17.835 seconds)

```

```

select min(subway) from mta;
0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select min(subway) from mta;
INFO : Compiling command(queryId=hive_20220426170935_37fe9c30-1e89-4318-8f62-b188d1f10940): se
lect min(subway) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:int, comment:nul
l)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426170935_37fe9c30-1e89-4318-8f62-b188d1f
10940); Time taken: 54.604 seconds
INFO : Executing command(queryId=hive_20220426170935_37fe9c30-1e89-4318-8f62-b188d1f10940): se
lect min(subway) from mta
WARN :
INFO : Query ID = hive_20220426170935_37fe9c30-1e89-4318-8f62-b188d1f10940
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33606
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33606/
INFO : Starting Job = job_1648648882306_33606, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33606/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33606
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 17:10:36,627 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 17:10:41,699 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.92 sec
INFO : 2022-04-26 17:10:46,764 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 4.83 sec
INFO : MapReduce Total cumulative CPU time: 4 seconds 830 msec
INFO : Ended Job = job_1648648882306_33606
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.83 sec HDFS Read: 72818 HDFS Wri
te: 106 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 4 seconds 830 msec
INFO : Completed executing command(queryId=hive_20220426170935_37fe9c30-1e89-4318-8f62-b188d1f
10940); Time taken: 18.303 seconds
INFO : OK
+-----+
| _c0 |
+-----+
| 198693 |
+-----+
1 row selected (72.92 seconds)

```

```

select avg(subway) from mta;
0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select avg(subway) from mta;
INFO : Compiling command(queryId=hive_20220426171118_9b349840-c040-4e45-b1bb-eca51671c47b): se
lect avg(subway) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426171118_9b349840-c040-4e45-b1bb-eca5167
1c47b); Time taken: 0.102 seconds
INFO : Executing command(queryId=hive_20220426171118_9b349840-c040-4e45-b1bb-eca51671c47b): se
lect avg(subway) from mta
WARN :
INFO : Query ID = hive_20220426171118_9b349840-c040-4e45-b1bb-eca51671c47b
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33607
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33607/
INFO : Starting Job = job_1648648882306_33607, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33607/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33607
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 17:11:23,968 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 17:11:30,048 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.93 sec
INFO : 2022-04-26 17:11:35,115 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 4.82 sec
INFO : MapReduce Total cumulative CPU time: 4 seconds 820 msec
INFO : Ended Job = job_1648648882306_33607
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 4.82 sec   HDFS Read: 73065 HDFS Wri
te: 118 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 4 seconds 820 msec
INFO : Completed executing command(queryId=hive_20220426171118_9b349840-c040-4e45-b1bb-eca5167
1c47b); Time taken: 17.468 seconds
INFO : OK
+-----+
|      _c0      |
+-----+
| 1771166.2398373983  |
+-----+
1 row selected (17.586 seconds)

```

```
select variance(subway) from mta;
```

```
0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select variance(subway) from mta;
INFO  : Compiling command(queryId=hive_20220426171229_e9c13080-eb84-4751-9790-c5fde29e077f): se
lect variance(subway) from mta
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20220426171229_e9c13080-eb84-4751-9790-c5fde29
e077f); Time taken: 0.109 seconds
INFO  : Executing command(queryId=hive_20220426171229_e9c13080-eb84-4751-9790-c5fde29e077f): se
lect variance(subway) from mta
WARN  :
INFO  : Query ID = hive_20220426171229_e9c13080-eb84-4751-9790-c5fde29e077f
INFO  : Total jobs = 1
INFO  : Launching Job 1 out of 1
INFO  : Starting task [Stage-1:MAPRED] in serial mode
INFO  : Number of reduce tasks determined at compile time: 1
INFO  : In order to change the average load for a reducer (in bytes):
INFO  :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO  : In order to limit the maximum number of reducers:
INFO  :   set hive.exec.reducers.max=<number>
INFO  : In order to set a constant number of reducers:
INFO  :   set mapreduce.job.reduces=<number>
INFO  : number of splits:1
INFO  : Submitting tokens for job: job_1648648882306_33610
INFO  : Executing with tokens: []
INFO  : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33610/
INFO  : Starting Job = job_1648648882306_33610, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33610/
INFO  : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33610
INFO  : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO  : 2022-04-26 17:12:35,162 Stage-1 map = 0%,  reduce = 0%
INFO  : 2022-04-26 17:12:41,239 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.96 sec
INFO  : 2022-04-26 17:12:47,313 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 5.93 sec
INFO  : MapReduce Total cumulative CPU time: 5 seconds 930 msec
INFO  : Ended Job = job_1648648882306_33610
INFO  : MapReduce Jobs Launched:
INFO  : Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 5.93 sec  HDFS Read: 73079 HDFS Wri
te: 120 HDFS EC Read: 0 SUCCESS
INFO  : Total MapReduce CPU Time Spent: 5 seconds 930 msec
INFO  : Completed executing command(queryId=hive_20220426171229_e9c13080-eb84-4751-9790-c5fde29
e077f); Time taken: 18.56 seconds
INFO  : OK
+-----+
|      _c0      |
+-----+
| 6.954712613195287E11  |
+-----+
1 row selected (18.682 seconds)
```

```

select stddev_pop(subway) from mta;
0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select stddev_pop(subway) from mta;
INFO : Compiling command(queryId=hive_20220426171315_29177d94-c29f-4b76-ab30-08c38151b372): se
lect stddev_pop(subway) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426171315_29177d94-c29f-4b76-ab30-08c3815
1b372); Time taken: 0.133 seconds
INFO : Executing command(queryId=hive_20220426171315_29177d94-c29f-4b76-ab30-08c38151b372): se
lect stddev_pop(subway) from mta
WARN :
INFO : Query ID = hive_20220426171315_29177d94-c29f-4b76-ab30-08c38151b372
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33612
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33612/
INFO : Starting Job = job_1648648882306_33612, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33612/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33612
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 17:13:21,790 Stage-1 map = 0%, reduce = 0%
INFO : 2022-04-26 17:13:26,857 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.95 sec
INFO : 2022-04-26 17:13:32,929 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6.04 sec
INFO : MapReduce Total cumulative CPU time: 6 seconds 40 msec
INFO : Ended Job = job_1648648882306_33612
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.04 sec    HDFS Read: 73069 HDFS Wri
te: 117 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 40 msec
INFO : Completed executing command(queryId=hive_20220426171315_29177d94-c29f-4b76-ab30-08c3815
1b372); Time taken: 18.418 seconds
INFO : OK
+-----+
|      _c0      |
+-----+
| 833949.1958863733 |
+-----+
1 row selected (18.561 seconds)

```

```

select max(subway_percent) from mta;
```

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select max(subway_percent) from mta;
INFO : Compiling command(queryId=hive_20220426171401_4632fb20-2d6a-4169-97e0-e764efba8052): se
lect max(subway_percent) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426171401_4632fb20-2d6a-4169-97e0-e764efb
a8052); Time taken: 0.095 seconds
INFO : Executing command(queryId=hive_20220426171401_4632fb20-2d6a-4169-97e0-e764efba8052): se
lect max(subway_percent) from mta
WARN :
INFO : Query ID = hive_20220426171401_4632fb20-2d6a-4169-97e0-e764efba8052
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33615
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33615/
INFO : Starting Job = job_1648648882306_33615, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33615/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33615
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 17:14:06,215 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 17:14:12,295 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.84 sec
INFO : 2022-04-26 17:14:18,366 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 6.26 sec
INFO : MapReduce Total cumulative CPU time: 6 seconds 260 msec
INFO : Ended Job = job_1648648882306_33615
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.26 sec HDFS Read: 72836 HDFS Wri
te: 105 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 260 msec
INFO : Completed executing command(queryId=hive_20220426171401_4632fb20-2d6a-4169-97e0-e764efb
a8052); Time taken: 18.198 seconds
INFO : OK
+---+
| _c0 |
+---+
| 0.729 |
+---+
1 row selected (18.304 seconds)

```

select min(subway_percent) from mta;

```
[0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select min(subway_percent) from mta;
INFO  : Compiling command(queryId=hive_20220426171447_6142c735-f557-4636-87a8-fc3c0bc0e444): se
lect min(subway_percent) from mta
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20220426171447_6142c735-f557-4636-87a8-fc3c0bc
0e444); Time taken: 0.096 seconds
INFO  : Executing command(queryId=hive_20220426171447_6142c735-f557-4636-87a8-fc3c0bc0e444): se
lect min(subway_percent) from mta
WARN  :
INFO  : Query ID = hive_20220426171447_6142c735-f557-4636-87a8-fc3c0bc0e444
INFO  : Total jobs = 1
INFO  : Launching Job 1 out of 1
INFO  : Starting task [Stage-1:MAPRED] in serial mode
INFO  : Number of reduce tasks determined at compile time: 1
INFO  : In order to change the average load for a reducer (in bytes):
INFO  :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO  : In order to limit the maximum number of reducers:
INFO  :   set hive.exec.reducers.max=<number>
INFO  : In order to set a constant number of reducers:
INFO  :   set mapreduce.job.reduces=<number>
INFO  : number of splits:1
INFO  : Submitting tokens for job: job_1648648882306_33620
INFO  : Executing with tokens: []
INFO  : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33620/
INFO  : Starting Job = job_1648648882306_33620, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33620/
INFO  : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33620
INFO  : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO  : 2022-04-26 17:14:52,354 Stage-1 map = 0%,  reduce = 0%
INFO  : 2022-04-26 17:14:58,431 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 2.06 sec
INFO  : 2022-04-26 17:15:05,521 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 4.7 sec
INFO  : MapReduce Total cumulative CPU time: 4 seconds 700 msec
INFO  : Ended Job = job_1648648882306_33620
INFO  : MapReduce Jobs Launched:
INFO  : Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 4.7 sec  HDFS Read: 72836 HDFS Writ
e: 105 HDFS EC Read: 0 SUCCESS
INFO  : Total MapReduce CPU Time Spent: 4 seconds 700 msec
INFO  : Completed executing command(queryId=hive_20220426171447_6142c735-f557-4636-87a8-fc3c0bc
0e444); Time taken: 19.247 seconds
INFO  : OK
+-----+
| _c0 |
+-----+
| 0.065 |
+-----+
1 row selected (19.355 seconds)
```

select avg(subway_percent) from mta;

```

[0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select avg(subway_percent) from mta;
INFO  : Compiling command(queryId=hive_20220426171538_592d5500-7a33-4c64-9543-d0e9704527a6): se
lect avg(subway_percent) from mta
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20220426171538_592d5500-7a33-4c64-9543-d0e9704
527a6); Time taken: 0.112 seconds
INFO  : Executing command(queryId=hive_20220426171538_592d5500-7a33-4c64-9543-d0e9704527a6): se
lect avg(subway_percent) from mta
WARN  :
INFO  : Query ID = hive_20220426171538_592d5500-7a33-4c64-9543-d0e9704527a6
INFO  : Total jobs = 1
INFO  : Launching Job 1 out of 1
INFO  : Starting task [Stage-1:MAPRED] in serial mode
INFO  : Number of reduce tasks determined at compile time: 1
INFO  : In order to change the average load for a reducer (in bytes):
INFO  :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO  : In order to limit the maximum number of reducers:
INFO  :   set hive.exec.reducers.max=<number>
INFO  : In order to set a constant number of reducers:
INFO  :   set mapreduce.job.reduces=<number>
INFO  : number of splits:1
INFO  : Submitting tokens for job: job_1648648882306_33621
INFO  : Executing with tokens: []
INFO  : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33621/
INFO  : Starting Job = job_1648648882306_33621, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33621/
INFO  : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33621
INFO  : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO  : 2022-04-26 17:15:44,980 Stage-1 map = 0%,  reduce = 0%
INFO  : 2022-04-26 17:15:50,052 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 2.23 sec
INFO  : 2022-04-26 17:15:56,125 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 7.06 sec
INFO  : MapReduce Total cumulative CPU time: 7 seconds 60 msec
INFO  : Ended Job = job_1648648882306_33621
INFO  : MapReduce Jobs Launched:
INFO  : Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 7.06 sec  HDFS Read: 73100 HDFS Wri
te: 119 HDFS EC Read: 0 SUCCESS
INFO  : Total MapReduce CPU Time Spent: 7 seconds 60 msec
INFO  : Completed executing command(queryId=hive_20220426171538_592d5500-7a33-4c64-9543-d0e9704
527a6); Time taken: 18.681 seconds
INFO  : OK
+-----+
|      _c0      |
+-----+
| 0.38512195121951215 |
+-----+
1 row selected (18.804 seconds)

```

select variance(subway_percent) from mta;

```

[0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select variance(subway_percent) from mta;
INFO : Compiling command(queryId=hive_20220426171615_0951289b-bf8c-466d-8876-3e478c9c855d): se
lect variance(subway_percent) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426171615_0951289b-bf8c-466d-8876-3e478c9
c855d); Time taken: 0.126 seconds
INFO : Executing command(queryId=hive_20220426171615_0951289b-bf8c-466d-8876-3e478c9c855d): se
lect variance(subway_percent) from mta
WARN :
INFO : Query ID = hive_20220426171615_0951289b-bf8c-466d-8876-3e478c9c855d
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33622
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33622/
INFO : Starting Job = job_1648648882306_33622, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33622/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33622
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 17:16:20,254 Stage-1 map = 0%, reduce = 0%
INFO : 2022-04-26 17:16:26,331 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.98 sec
INFO : 2022-04-26 17:16:31,388 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6.57 sec
INFO : MapReduce Total cumulative CPU time: 6 seconds 570 msec
INFO : Ended Job = job_1648648882306_33622
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.57 sec HDFS Read: 73108 HDFS Wri
te: 120 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 570 msec
INFO : Completed executing command(queryId=hive_20220426171615_0951289b-bf8c-466d-8876-3e478c9
c855d); Time taken: 17.298 seconds
INFO : OK
+-----+
|      _c0      |
+-----+
| 0.024114012228171063 |
+-----+
1 row selected (17.436 seconds)
2. jdbc:hive2://hm-1.hpc.nyu.edu:10000/>
```

select stddev_pop(subway_percent) from mta;

```

[0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select stddev_pop(subway_percent) from mta;
INFO : Compiling command(queryId=hive_20220426171734_d8f96aeb-ab06-4f44-b77c-cf022f91be02): se
lect stddev_pop(subway_percent) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426171734_d8f96aeb-ab06-4f44-b77c-cf022f9
1be02); Time taken: 0.092 seconds
INFO : Executing command(queryId=hive_20220426171734_d8f96aeb-ab06-4f44-b77c-cf022f91be02): se
lect stddev_pop(subway_percent) from mta
WARN :
INFO : Query ID = hive_20220426171734_d8f96aeb-ab06-4f44-b77c-cf022f91be02
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33630
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33630/
INFO : Starting Job = job_1648648882306_33630, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33630/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33630
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 17:17:39,266 Stage-1 map = 0%, reduce = 0%
INFO : 2022-04-26 17:17:44,342 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.86 sec
INFO : 2022-04-26 17:17:50,421 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6.78 sec
INFO : MapReduce Total cumulative CPU time: 6 seconds 780 msec
INFO : Ended Job = job_1648648882306_33630
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.78 sec HDFS Read: 73008 HDFS Wri
te: 118 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 780 msec
INFO : Completed executing command(queryId=hive_20220426171734_d8f96aeb-ab06-4f44-b77c-cf022f9
1be02); Time taken: 17.274 seconds
INFO : OK
+-----+
|      _c0      |
+-----+
| 0.1552868707527171 |
+-----+
1 row selected (17.379 seconds)
a. jdbc:hive2://hm-1.hpc.nyu.edu:10000/ ~
```

select max(ride) from mta;

```

[0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select max(ride) from mta;
INFO  : Compiling command(queryId=hive_20220426171837_fe43182a-e6bb-4654-a8dc-62a578616374): se
lect max(ride) from mta
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:int, comment:nul
l)], properties:null)
INFO  : Completed compiling command(queryId=hive_20220426171837_fe43182a-e6bb-4654-a8dc-62a5786
16374); Time taken: 0.125 seconds
INFO  : Executing command(queryId=hive_20220426171837_fe43182a-e6bb-4654-a8dc-62a578616374): se
lect max(ride) from mta
WARN  :
INFO  : Query ID = hive_20220426171837_fe43182a-e6bb-4654-a8dc-62a578616374
INFO  : Total jobs = 1
INFO  : Launching Job 1 out of 1
INFO  : Starting task [Stage-1:MAPRED] in serial mode
INFO  : Number of reduce tasks determined at compile time: 1
INFO  : In order to change the average load for a reducer (in bytes):
INFO  :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO  : In order to limit the maximum number of reducers:
INFO  :   set hive.exec.reducers.max=<number>
INFO  : In order to set a constant number of reducers:
INFO  :   set mapreduce.job.reduces=<number>
INFO  : number of splits:1
INFO  : Submitting tokens for job: job_1648648882306_33635
INFO  : Executing with tokens: []
INFO  : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33635/
INFO  : Starting Job = job_1648648882306_33635, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33635/
INFO  : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33635
INFO  : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO  : 2022-04-26 17:18:44,580 Stage-1 map = 0%,  reduce = 0%
INFO  : 2022-04-26 17:18:49,653 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.93 sec
INFO  : 2022-04-26 17:18:55,730 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 4.93 sec
INFO  : MapReduce Total cumulative CPU time: 4 seconds 930 msec
INFO  : Ended Job = job_1648648882306_33635
INFO  : MapReduce Jobs Launched:
INFO  : Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 4.93 sec  HDFS Read: 72814 HDFS Wri
te: 107 HDFS EC Read: 0 SUCCESS
INFO  : Total MapReduce CPU Time Spent: 4 seconds 930 msec
INFO  : Completed executing command(queryId=hive_20220426171837_fe43182a-e6bb-4654-a8dc-62a5786
16374); Time taken: 18.814 seconds
INFO  : OK
+-----+
| _c0  |
+-----+
| 1532668 |
+-----+
1 row selected (18.95 seconds)

```

select min(ride) from mta;

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select min(ride) from mta;
INFO : Compiling command(queryId=hive_20220426172009_df2edab3-f6a8-4df6-b4da-49f9320f7c2f): se
lect min(ride) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:int, comment:nul
l), properties:null])
INFO : Completed compiling command(queryId=hive_20220426172009_df2edab3-f6a8-4df6-b4da-49f9320
f7c2f); Time taken: 0.109 seconds
INFO : Executing command(queryId=hive_20220426172009_df2edab3-f6a8-4df6-b4da-49f9320f7c2f): se
lect min(ride) from mta
WARN :
INFO : Query ID = hive_20220426172009_df2edab3-f6a8-4df6-b4da-49f9320f7c2f
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33639
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33639/
INFO : Starting Job = job_1648648882306_33639, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33639/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33639
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 17:20:16,423 Stage-1 map = 0%, reduce = 0%
INFO : 2022-04-26 17:20:21,495 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.97 sec
INFO : 2022-04-26 17:20:27,606 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 4.85 sec
INFO : MapReduce Total cumulative CPU time: 4 seconds 850 msec
INFO : Ended Job = job_1648648882306_33639
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.85 sec HDFS Read: 72814 HDFS Wri
te: 106 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 4 seconds 850 msec
INFO : Completed executing command(queryId=hive_20220426172009_df2edab3-f6a8-4df6-b4da-49f9320
f7c2f); Time taken: 19.688 seconds
INFO : OK
+---+
| _c0 |
+---+
| 271959 |
+---+
1 row selected (19.808 seconds)

```

select avg(ride) from mta;

```
[0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select avg(ride) from mta;
INFO  : Compiling command(queryId=hive_20220426172112_01fec8f9-a3d5-448d-9753-ef6573eb26f4): se
lect avg(ride) from mta
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20220426172112_01fec8f9-a3d5-448d-9753-ef6573e
b26f4); Time taken: 0.11 seconds
INFO  : Executing command(queryId=hive_20220426172112_01fec8f9-a3d5-448d-9753-ef6573eb26f4): se
lect avg(ride) from mta
WARN  :
INFO  : Query ID = hive_20220426172112_01fec8f9-a3d5-448d-9753-ef6573eb26f4
INFO  : Total jobs = 1
INFO  : Launching Job 1 out of 1
INFO  : Starting task [Stage-1:MAPRED] in serial mode
INFO  : Number of reduce tasks determined at compile time: 1
INFO  : In order to change the average load for a reducer (in bytes):
INFO  :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO  : In order to limit the maximum number of reducers:
INFO  :   set hive.exec.reducers.max=<number>
INFO  : In order to set a constant number of reducers:
INFO  :   set mapreduce.job.reduces=<number>
INFO  : number of splits:1
INFO  : Submitting tokens for job: job_1648648882306_33642
INFO  : Executing with tokens: []
INFO  : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33642/
INFO  : Starting Job = job_1648648882306_33642, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33642/
INFO  : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33642
INFO  : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO  : 2022-04-26 17:21:18,799 Stage-1 map = 0%,  reduce = 0%
INFO  : 2022-04-26 17:21:23,867 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 2.19 sec
INFO  : 2022-04-26 17:21:29,938 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 5.09 sec
INFO  : MapReduce Total cumulative CPU time: 5 seconds 90 msec
INFO  : Ended Job = job_1648648882306_33642
INFO  : MapReduce Jobs Launched:
INFO  : Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 5.09 sec  HDFS Read: 73061 HDFS Wri
te: 117 HDFS EC Read: 0 SUCCESS
INFO  : Total MapReduce CPU Time Spent: 5 seconds 90 msec
INFO  : Completed executing command(queryId=hive_20220426172112_01fec8f9-a3d5-448d-9753-ef6573e
b26f4); Time taken: 18.664 seconds
INFO  : OK
+-----+
|      _c0      |
+-----+
| 984293.7046070461 |
+-----+
1 row selected (18.786 seconds)
```

select variance(ride) from mta;

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select variance(ride) from mta; ]
INFO : Compiling command(queryId=hive_20220426172152_e5bbc3bd-da37-44d2-8367-41f1573d3b2a): se
lect variance(ride) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426172152_e5bbc3bd-da37-44d2-8367-41f1573
d3b2a); Time taken: 0.125 seconds
INFO : Executing command(queryId=hive_20220426172152_e5bbc3bd-da37-44d2-8367-41f1573d3b2a): se
lect variance(ride) from mta
WARN :
INFO : Query ID = hive_20220426172152_e5bbc3bd-da37-44d2-8367-41f1573d3b2a
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33644
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33644/
INFO : Starting Job = job_1648648882306_33644, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33644/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33644
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 17:21:58,261 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 17:22:04,338 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.96 sec
INFO : 2022-04-26 17:22:10,409 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 6.34 sec
INFO : MapReduce Total cumulative CPU time: 6 seconds 340 msec
INFO : Ended Job = job_1648648882306_33644
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 6.34 sec  HDFS Read: 73075 HDFS Wri
te: 120 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 340 msec
INFO : Completed executing command(queryId=hive_20220426172152_e5bbc3bd-da37-44d2-8367-41f1573
d3b2a); Time taken: 18.461 seconds
INFO : OK
+-----+
|      _c0      |
+-----+
| 9.949155225650896E10 |
+-----+
1 row selected (18.598 seconds)

```

select stddev_pop(ride) from mta;

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select stddev_pop(ride) from mta;
INFO : Compiling command(queryId=hive_20220426172235_be9134d8-7255-45c4-9af3-1717f86788e4): se
lect stddev_pop(ride) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426172235_be9134d8-7255-45c4-9af3-1717f86
788e4); Time taken: 0.117 seconds
INFO : Executing command(queryId=hive_20220426172235_be9134d8-7255-45c4-9af3-1717f86788e4): se
lect stddev_pop(ride) from mta
WARN :
INFO : Query ID = hive_20220426172235_be9134d8-7255-45c4-9af3-1717f86788e4
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33647
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33647/
INFO : Starting Job = job_1648648882306_33647, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33647/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33647
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 17:22:40,691 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 17:22:46,773 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.96 sec
INFO : 2022-04-26 17:22:51,839 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 6.41 sec
INFO : MapReduce Total cumulative CPU time: 6 seconds 410 msec
INFO : Ended Job = job_1648648882306_33647
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.41 sec HDFS Read: 73079 HDFS Wri
te: 117 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 410 msec
INFO : Completed executing command(queryId=hive_20220426172235_be9134d8-7255-45c4-9af3-1717f86
788e4); Time taken: 17.297 seconds
INFO : OK
+-----+
|      _c0      |
+-----+
| 315422.8150538717 |
+-----+
1 row selected (17.427 seconds)

```

select min(ride_percent) from mta;

```
[0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select min(ride_percent) from mta;
INFO  : Compiling command(queryId=hive_20220426172340_00290bc6-8005-40e2-b478-4dc9c66ff738): se
lect min(ride_percent) from mta
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20220426172340_00290bc6-8005-40e2-b478-4dc9c66
ff738); Time taken: 0.105 seconds
INFO  : Executing command(queryId=hive_20220426172340_00290bc6-8005-40e2-b478-4dc9c66ff738): se
lect min(ride_percent) from mta
WARN  :
INFO  : Query ID = hive_20220426172340_00290bc6-8005-40e2-b478-4dc9c66ff738
INFO  : Total jobs = 1
INFO  : Launching Job 1 out of 1
INFO  : Starting task [Stage-1:MAPRED] in serial mode
INFO  : Number of reduce tasks determined at compile time: 1
INFO  : In order to change the average load for a reducer (in bytes):
INFO  :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO  : In order to limit the maximum number of reducers:
INFO  :   set hive.exec.reducers.max=<number>
INFO  : In order to set a constant number of reducers:
INFO  :   set mapreduce.job.reduces=<number>
INFO  : number of splits:1
INFO  : Submitting tokens for job: job_1648648882306_33649
INFO  : Executing with tokens: []
INFO  : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33649/
INFO  : Starting Job = job_1648648882306_33649, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33649/
INFO  : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33649
INFO  : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO  : 2022-04-26 17:23:47,088 Stage-1 map = 0%,  reduce = 0%
INFO  : 2022-04-26 17:23:52,166 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.93 sec
INFO  : 2022-04-26 17:23:58,257 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 4.81 sec
INFO  : MapReduce Total cumulative CPU time: 4 seconds 810 msec
INFO  : Ended Job = job_1648648882306_33649
INFO  : MapReduce Jobs Launched:
INFO  : Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 4.81 sec  HDFS Read: 72728 HDFS Wri
te: 105 HDFS EC Read: 0 SUCCESS
INFO  : Total MapReduce CPU Time Spent: 4 seconds 810 msec
INFO  : Completed executing command(queryId=hive_20220426172340_00290bc6-8005-40e2-b478-4dc9c66
ff738); Time taken: 18.247 seconds
INFO  : OK
+-----+
| _c0 |
+-----+
| 0.154 |
+-----+
1 row selected (18.364 seconds)
```

select max(ride_percent) from mta;

```

[0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select max(ride_percent) from mta;
INFO : Compiling command(queryId=hive_20220426172510_0572a914-a077-40d9-9e82-9c5e7fe57043): se
lect max(ride_percent) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426172510_0572a914-a077-40d9-9e82-9c5e7fe
57043); Time taken: 0.123 seconds
INFO : Executing command(queryId=hive_20220426172510_0572a914-a077-40d9-9e82-9c5e7fe57043): se
lect max(ride_percent) from mta
WARN :
INFO : Query ID = hive_20220426172510_0572a914-a077-40d9-9e82-9c5e7fe57043
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33651
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33651/
INFO : Starting Job = job_1648648882306_33651, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33651/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33651
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 17:25:15,894 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 17:25:21,979 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.94 sec
INFO : 2022-04-26 17:25:28,057 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 6.65 sec
INFO : MapReduce Total cumulative CPU time: 6 seconds 650 msec
INFO : Ended Job = job_1648648882306_33651
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.65 sec HDFS Read: 72832 HDFS Wri
te: 104 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 650 msec
INFO : Completed executing command(queryId=hive_20220426172510_0572a914-a077-40d9-9e82-9c5e7fe
57043); Time taken: 18.285 seconds
INFO : OK
+-----+
| _c0 |
+-----+
| 0.87 |
+-----+
1 row selected (18.424 seconds)

```

select avg(ride_percent) from mta;

```
[0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select avg(ride_percent) from mta;
INFO  : Compiling command(queryId=hive_20220426172611_295eaf85-ec51-48e3-aff7-678637444eb7): se
lect avg(ride_percent) from mta
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20220426172611_295eaf85-ec51-48e3-aff7-6786374
4eb7); Time taken: 0.1 seconds
INFO  : Executing command(queryId=hive_20220426172611_295eaf85-ec51-48e3-aff7-678637444eb7): se
lect avg(ride_percent) from mta
WARN  :
INFO  : Query ID = hive_20220426172611_295eaf85-ec51-48e3-aff7-678637444eb7
INFO  : Total jobs = 1
INFO  : Launching Job 1 out of 1
INFO  : Starting task [Stage-1:MAPRED] in serial mode
INFO  : Number of reduce tasks determined at compile time: 1
INFO  : In order to change the average load for a reducer (in bytes):
INFO  :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO  : In order to limit the maximum number of reducers:
INFO  :   set hive.exec.reducers.max=<number>
INFO  : In order to set a constant number of reducers:
INFO  :   set mapreduce.job.reduces=<number>
INFO  : number of splits:1
INFO  : Submitting tokens for job: job_1648648882306_33653
INFO  : Executing with tokens: []
INFO  : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33653/
INFO  : Starting Job = job_1648648882306_33653, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33653/
INFO  : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33653
INFO  : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO  : 2022-04-26 17:26:17,283 Stage-1 map = 0%, reduce = 0%
INFO  : 2022-04-26 17:26:22,404 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.97 sec
INFO  : 2022-04-26 17:26:27,463 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6.86 sec
INFO  : MapReduce Total cumulative CPU time: 6 seconds 860 msec
INFO  : Ended Job = job_1648648882306_33653
INFO  : MapReduce Jobs Launched:
INFO  : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.86 sec HDFS Read: 73096 HDFS Wri
te: 118 HDFS EC Read: 0 SUCCESS
INFO  : Total MapReduce CPU Time Spent: 6 seconds 860 msec
INFO  : Completed executing command(queryId=hive_20220426172611_295eaf85-ec51-48e3-aff7-6786374
4eb7); Time taken: 17.24 seconds
INFO  : OK
+-----+
|      _c0      |
+-----+
| 0.5336504065040646 |
+-----+
1 row selected (17.352 seconds)
```

select variance(ride_percent) from mta;

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000/> select variance(ride_percent) from mta;
INFO : Compiling command(queryId=hive_20220426172720_23568a31-f0a6-4df6-b436-451c7b78b6d1): se
lect variance(ride_percent) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426172720_23568a31-f0a6-4df6-b436-451c7b7
8b6d1); Time taken: 0.097 seconds
INFO : Executing command(queryId=hive_20220426172720_23568a31-f0a6-4df6-b436-451c7b78b6d1): se
lect variance(ride_percent) from mta
WARN :
INFO : Query ID = hive_20220426172720_23568a31-f0a6-4df6-b436-451c7b78b6d1
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33656
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33656/
INFO : Starting Job = job_1648648882306_33656, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33656/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33656
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 17:27:26,090 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 17:27:32,170 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 2.05 sec
INFO : 2022-04-26 17:27:38,244 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 4.94 sec
INFO : MapReduce Total cumulative CPU time: 4 seconds 940 msec
INFO : Ended Job = job_1648648882306_33656
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.94 sec HDFS Read: 73104 HDFS Wri
te: 120 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 4 seconds 940 msec
INFO : Completed executing command(queryId=hive_20220426172720_23568a31-f0a6-4df6-b436-451c7b7
8b6d1); Time taken: 18.272 seconds
INFO : OK
+-----+
|      _c0      |
+-----+
| 0.012455495670566445 |
+-----+
1 row selected (18.381 seconds)

```

select stddev_pop(ride_percent) from mta;

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select stddev_pop(ride_percent) from mta;
INFO : Compiling command(queryId=hive_20220426172803_46dc58d9-da05-41db-b22a-250466e90ca2): se
lect stddev_pop(ride_percent) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426172803_46dc58d9-da05-41db-b22a-250466e
90ca2); Time taken: 0.078 seconds
INFO : Executing command(queryId=hive_20220426172803_46dc58d9-da05-41db-b22a-250466e90ca2): se
lect stddev_pop(ride_percent) from mta
WARN :
INFO : Query ID = hive_20220426172803_46dc58d9-da05-41db-b22a-250466e90ca2
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33657
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33657/
INFO : Starting Job = job_1648648882306_33657, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33657/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33657
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 17:28:08,453 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 17:28:14,547 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.84 sec
INFO : 2022-04-26 17:28:20,617 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 4.89 sec
INFO : MapReduce Total cumulative CPU time: 4 seconds 890 msec
INFO : Ended Job = job_1648648882306_33657
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.89 sec HDFS Read: 73108 HDFS Wri
te: 119 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 4 seconds 890 msec
INFO : Completed executing command(queryId=hive_20220426172803_46dc58d9-da05-41db-b22a-250466e
90ca2); Time taken: 18.273 seconds
INFO : OK
+-----+
|      _c0      |
+-----+
| 0.11160419199369907 |
+-----+
1 row selected (18.362 seconds)

```

select min(lirr) from mta;

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000/> select min(lirr) from mta; ]
INFO : Compiling command(queryId=hive_20220426173151_b9ea0985-164f-40e9-851d-3d7efe554931): se
lect min(lirr) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:int, comment:nul
l)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426173151_b9ea0985-164f-40e9-851d-3d7efe5
54931); Time taken: 0.094 seconds
INFO : Executing command(queryId=hive_20220426173151_b9ea0985-164f-40e9-851d-3d7efe554931): se
lect min(lirr) from mta
WARN :
INFO : Query ID = hive_20220426173151_b9ea0985-164f-40e9-851d-3d7efe554931
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33665
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33665/
INFO : Starting Job = job_1648648882306_33665, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33665/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33665
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 17:31:57,552 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 17:32:02,677 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 2.01 sec
INFO : 2022-04-26 17:32:08,745 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 4.76 sec
INFO : MapReduce Total cumulative CPU time: 4 seconds 760 msec
INFO : Ended Job = job_1648648882306_33665
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.76 sec HDFS Read: 72800 HDFS Wri
te: 104 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 4 seconds 760 msec
INFO : Completed executing command(queryId=hive_20220426173151_b9ea0985-164f-40e9-851d-3d7efe5
54931); Time taken: 18.231 seconds
INFO : OK
+---+
| _c0 |
+---+
| 1900 |
+---+
1 row selected (18.335 seconds)
select max(lirr) from mta;

```

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select max(lirr) from mta;
INFO : Compiling command(queryId=hive_20220426173326_131ae0c9-82c9-4ae4-b0d7-3e6c3cc67ad3): se
lect max(lirr) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:int, comment:nul
l)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426173326_131ae0c9-82c9-4ae4-b0d7-3e6c3cc
67ad3); Time taken: 0.077 seconds
INFO : Executing command(queryId=hive_20220426173326_131ae0c9-82c9-4ae4-b0d7-3e6c3cc67ad3): se
lect max(lirr) from mta
WARN :
INFO : Query ID = hive_20220426173326_131ae0c9-82c9-4ae4-b0d7-3e6c3cc67ad3
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33671
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33671/
INFO : Starting Job = job_1648648882306_33671, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33671/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33671
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 17:33:31,391 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 17:33:37,467 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.92 sec
INFO : 2022-04-26 17:33:43,536 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 5.78 sec
INFO : MapReduce Total cumulative CPU time: 5 seconds 780 msec
INFO : Ended Job = job_1648648882306_33671
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 5.78 sec  HDFS Read: 72710 HDFS Wri
te: 106 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 5 seconds 780 msec
INFO : Completed executing command(queryId=hive_20220426173326_131ae0c9-82c9-4ae4-b0d7-3e6c3cc
67ad3); Time taken: 18.455 seconds
INFO : OK
+-----+
| _c0 |
+-----+
| 172300 |
+-----+
1 row selected (18.543 seconds)

select avg(lirr) from mta;

```

```
0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select avg(lirr) from mta;
INFO  : Compiling command(queryId=hive_20220426173419_4d10dd51-9698-4782-921b-4c9f0e8389c9): se
lect avg(lirr) from mta
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20220426173419_4d10dd51-9698-4782-921b-4c9f0e8
389c9); Time taken: 0.12 seconds
INFO  : Executing command(queryId=hive_20220426173419_4d10dd51-9698-4782-921b-4c9f0e8389c9): se
lect avg(lirr) from mta
WARN  :
INFO  : Query ID = hive_20220426173419_4d10dd51-9698-4782-921b-4c9f0e8389c9
INFO  : Total jobs = 1
INFO  : Launching Job 1 out of 1
INFO  : Starting task [Stage-1:MAPRED] in serial mode
INFO  : Number of reduce tasks determined at compile time: 1
INFO  : In order to change the average load for a reducer (in bytes):
INFO  :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO  : In order to limit the maximum number of reducers:
INFO  :   set hive.exec.reducers.max=<number>
INFO  : In order to set a constant number of reducers:
INFO  :   set mapreduce.job.reduces=<number>
INFO  : number of splits:1
INFO  : Submitting tokens for job: job_1648648882306_33674
INFO  : Executing with tokens: []
INFO  : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33674/
INFO  : Starting Job = job_1648648882306_33674, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33674/
INFO  : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33674
INFO  : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO  : 2022-04-26 17:34:26,443 Stage-1 map = 0%,  reduce = 0%
INFO  : 2022-04-26 17:34:31,511 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.98 sec
INFO  : 2022-04-26 17:34:37,584 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 5.62 sec
INFO  : MapReduce Total cumulative CPU time: 5 seconds 620 msec
INFO  : Ended Job = job_1648648882306_33674
INFO  : MapReduce Jobs Launched:
INFO  : Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 5.62 sec  HDFS Read: 73061 HDFS Wri
te: 117 HDFS EC Read: 0 SUCCESS
INFO  : Total MapReduce CPU Time Spent: 5 seconds 620 msec
INFO  : Completed executing command(queryId=hive_20220426173419_4d10dd51-9698-4782-921b-4c9f0e8
389c9); Time taken: 18.773 seconds
INFO  : OK
+-----+
|      _c0      |
+-----+
| 82617.75067750677 |
+-----+
1 row selected (18.903 seconds)
2. idhc:hive2?://hm-1.hpc.nyu.edu:10000/>
```

```
0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select variance(lirr) from mta;
INFO : Compiling command(queryId=hive_20220426173513_1a37a5a4-c6d2-4e7d-bdfe-c95eff1b537): se
lect variance(lirr) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426173513_1a37a5a4-c6d2-4e7d-bdfe-c95eff
1b537); Time taken: 0.135 seconds
INFO : Executing command(queryId=hive_20220426173513_1a37a5a4-c6d2-4e7d-bdfe-c95eff1b537): se
lect variance(lirr) from mta
WARN :
INFO : Query ID = hive_20220426173513_1a37a5a4-c6d2-4e7d-bdfe-c95eff1b537
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :     set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :     set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :     set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33677
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33677/
INFO : Starting Job = job_1648648882306_33677, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33677/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33677
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 17:35:19,978 Stage-1 map = 0%, reduce = 0%
INFO : 2022-04-26 17:35:25,047 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.93 sec
INFO : 2022-04-26 17:35:31,118 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 4.99 sec
INFO : MapReduce Total cumulative CPU time: 4 seconds 990 msec
INFO : Ended Job = job_1648648882306_33677
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.99 sec HDFS Read: 73075 HDFS Wri
te: 120 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 4 seconds 990 msec
INFO : Completed executing command(queryId=hive_20220426173513_1a37a5a4-c6d2-4e7d-bdfe-c95eff
1b537); Time taken: 19.717 seconds
INFO : OK
+-----+
| _c0 |
+-----+
| 1.8708354437210352E9 |
+-----+
1 row selected (19.864 seconds)
select stddev_pop(lirr) from mta;
```

```
[0: jdbc:hive2://hm-1.hpc.nyu.edu:10000]> select stddev_pop(lirr) from mta;
INFO : Compiling command(queryId=hive_20220426173559_92ebc8bd-0e02-4795-98d2-7cfa58b0622a): se
lect stddev_pop(lirr) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426173559_92ebc8bd-0e02-4795-98d2-7cfa58b
0622a); Time taken: 0.116 seconds
INFO : Executing command(queryId=hive_20220426173559_92ebc8bd-0e02-4795-98d2-7cfa58b0622a): se
lect stddev_pop(lirr) from mta
WARN :
INFO : Query ID = hive_20220426173559_92ebc8bd-0e02-4795-98d2-7cfa58b0622a
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33678
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33678/
INFO : Starting Job = job_1648648882306_33678, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33678/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33678
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 17:37:56,845 Stage-1 map = 0%, reduce = 0%
INFO : 2022-04-26 17:38:02,922 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.04 sec
INFO : 2022-04-26 17:38:08,993 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 4.89 sec
INFO : MapReduce Total cumulative CPU time: 4 seconds 890 msec
INFO : Ended Job = job_1648648882306_33678
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.89 sec HDFS Read: 73079 HDFS Wri
te: 117 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 4 seconds 890 msec
INFO : Completed executing command(queryId=hive_20220426173559_92ebc8bd-0e02-4795-98d2-7cfa58b
0622a); Time taken: 130.762 seconds
INFO : OK
+-----+
| _c0 |
+-----+
| 43253.15530364271 |
+-----+
1 row selected (130.889 seconds)
select min(lirr_percent) from mta;
```

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000/> select min(lIRR_percent) from mta;
INFO : Compiling command(queryId=hive_20220426174413_fb351e95-f289-42bb-8333-ebfc601fa40b): se
lect min(lIRR_percent) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426174413_fb351e95-f289-42bb-8333-ebfc601
fa40b); Time taken: 0.104 seconds
INFO : Executing command(queryId=hive_20220426174413_fb351e95-f289-42bb-8333-ebfc601fa40b): se
lect min(lIRR_percent) from mta
WARN :
INFO : Query ID = hive_20220426174413_fb351e95-f289-42bb-8333-ebfc601fa40b
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33696
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33696/
INFO : Starting Job = job_1648648882306_33696, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33696/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33696
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 17:44:19,624 Stage-1 map = 0%, reduce = 0%
INFO : 2022-04-26 17:44:24,695 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.02 sec
INFO : 2022-04-26 17:44:30,769 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 5.0 sec
INFO : MapReduce Total cumulative CPU time: 5 seconds 0 msec
INFO : Ended Job = job_1648648882306_33696
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 5.0 sec    HDFS Read: 72727 HDFS Writ
e: 104 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 5 seconds 0 msec
INFO : Completed executing command(queryId=hive_20220426174413_fb351e95-f289-42bb-8333-ebfc601
fa40b); Time taken: 18.223 seconds
INFO : OK
+---+
| _c0 |
+---+
| 0.02 |
+---+
1 row selected (18.339 seconds)
0: jdbc:hive2://hm-1.hpc.nyu.edu:10000/>
```

select max(lIRR_percent) from mta;

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select max(lIRR_percent) from mta;
INFO : Compiling command(queryId=hive_20220426174600_c0b74641-67a7-4bd1-9f3d-55cae1f78272): se
lect max(lIRR_percent) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426174600_c0b74641-67a7-4bd1-9f3d-55cae1f
78272); Time taken: 0.095 seconds
INFO : Executing command(queryId=hive_20220426174600_c0b74641-67a7-4bd1-9f3d-55cae1f78272): se
lect max(lIRR_percent) from mta
WARN :
INFO : Query ID = hive_20220426174600_c0b74641-67a7-4bd1-9f3d-55cae1f78272
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33700
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33700/
INFO : Starting Job = job_1648648882306_33700, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33700/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33700
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 17:47:19,450 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 17:47:24,520 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.97 sec
INFO : 2022-04-26 17:47:30,590 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 4.78 sec
INFO : MapReduce Total cumulative CPU time: 4 seconds 780 msec
INFO : Ended Job = job_1648648882306_33700
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.78 sec HDFS Read: 72831 HDFS Wri
te: 104 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 4 seconds 780 msec
INFO : Completed executing command(queryId=hive_20220426174600_c0b74641-67a7-4bd1-9f3d-55cae1f
78272); Time taken: 91.035 seconds
INFO : OK
+---+
| _c0 |
+---+
| 0.87 |
+---+
1 row selected (91.141 seconds)

```

select avg(lIRR_percent) from mta;

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000/> select avg(lIRR_percent) from mta;
INFO : Compiling command(queryId=hive_20220426174857_e7beae17-bfa6-4964-a9ec-27135cc5af5a): se
lect avg(lIRR_percent) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426174857_e7beae17-bfa6-4964-a9ec-27135cc
5af5a); Time taken: 0.117 seconds
INFO : Executing command(queryId=hive_20220426174857_e7beae17-bfa6-4964-a9ec-27135cc5af5a): se
lect avg(lIRR_percent) from mta
WARN :
INFO : Query ID = hive_20220426174857_e7beae17-bfa6-4964-a9ec-27135cc5af5a
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33707
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33707/
INFO : Starting Job = job_1648648882306_33707, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33707/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33707
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 17:49:04,346 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 17:49:09,415 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.91 sec
INFO : 2022-04-26 17:49:14,477 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 7.5 sec
INFO : MapReduce Total cumulative CPU time: 7 seconds 500 msec
INFO : Ended Job = job_1648648882306_33707
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 7.5 sec    HDFS Read: 73090 HDFS Writ
e: 119 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 7 seconds 500 msec
INFO : Completed executing command(queryId=hive_20220426174857_e7beae17-bfa6-4964-a9ec-27135cc
5af5a); Time taken: 17.756 seconds
INFO : OK
+-----+
|      _c0      |
+-----+
| 0.35994579945799354 |
+-----+
1 row selected (17.884 seconds)

select variance(lIRR_percent) from mta;

```

```

[0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select variance(lIRR_percent) from mta;
INFO  : Compiling command(queryId=hive_20220426175151_855ed680-eeb3-4fc9-a080-fb79f7c98be0): se
lect variance(lIRR_percent) from mta
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20220426175151_855ed680-eeb3-4fc9-a080-fb79f7c
98be0); Time taken: 0.176 seconds
INFO  : Executing command(queryId=hive_20220426175151_855ed680-eeb3-4fc9-a080-fb79f7c98be0): se
lect variance(lIRR_percent) from mta
WARN  :
INFO  : Query ID = hive_20220426175151_855ed680-eeb3-4fc9-a080-fb79f7c98be0
INFO  : Total jobs = 1
INFO  : Launching Job 1 out of 1
INFO  : Starting task [Stage-1:MAPRED] in serial mode
INFO  : Number of reduce tasks determined at compile time: 1
INFO  : In order to change the average load for a reducer (in bytes):
INFO  :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO  : In order to limit the maximum number of reducers:
INFO  :   set hive.exec.reducers.max=<number>
INFO  : In order to set a constant number of reducers:
INFO  :   set mapreduce.job.reduces=<number>
INFO  : number of splits:1
INFO  : Submitting tokens for job: job_1648648882306_33716
INFO  : Executing with tokens: []
INFO  : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33716/
INFO  : Starting Job = job_1648648882306_33716, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33716/
INFO  : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33716
INFO  : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO  : 2022-04-26 17:51:58,145 Stage-1 map = 0%,  reduce = 0%
INFO  : 2022-04-26 17:52:03,214 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.9 sec
INFO  : 2022-04-26 17:52:08,274 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 6.6 sec
INFO  : MapReduce Total cumulative CPU time: 6 seconds 600 msec
INFO  : Ended Job = job_1648648882306_33716
INFO  : MapReduce Jobs Launched:
INFO  : Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 6.6 sec  HDFS Read: 73103 HDFS Writ
e: 119 HDFS EC Read: 0 SUCCESS
INFO  : Total MapReduce CPU Time Spent: 6 seconds 600 msec
INFO  : Completed executing command(queryId=hive_20220426175151_855ed680-eeb3-4fc9-a080-fb79f7c
98be0); Time taken: 18.93 seconds
INFO  : OK
+-----+
|      _c0      |
+-----+
| 0.03297587782110884 |
+-----+
1 row selected (19.117 seconds)

select stdDev_pop(lIRR_percent) from mta;
```

```

select min(metro_north) from mta;
0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select stddev_pop(lirr_percent) from mta;
INFO : Compiling command(queryId=hive_20220426175331_07731c09-25da-453e-854f-e494b9ecbb4a): se
lect stddev_pop(lirr_percent) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426175331_07731c09-25da-453e-854f-e494b9
ecbb4a); Time taken: 0.127 seconds
INFO : Executing command(queryId=hive_20220426175331_07731c09-25da-453e-854f-e494b9ecbb4a): se
lect stddev_pop(lirr_percent) from mta
WARN :
INFO : Query ID = hive_20220426175331_07731c09-25da-453e-854f-e494b9ecbb4a
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33722
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33722/
INFO : Starting Job = job_1648648882306_33722, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33722/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33722
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 17:53:37,408 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 17:53:43,490 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.97 sec
INFO : 2022-04-26 17:53:48,555 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 5.05 sec
INFO : MapReduce Total cumulative CPU time: 5 seconds 50 msec
INFO : Ended Job = job_1648648882306_33722
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 5.05 sec HDFS Read: 73005 HDFS Wri
te: 118 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 5 seconds 50 msec
INFO : Completed executing command(queryId=hive_20220426175331_07731c09-25da-453e-854f-e494b9
ecbb4a); Time taken: 17.675 seconds
INFO : OK
+-----+
|      _c0      |
+-----+
| 0.1815926149960643 |
+-----+
1 row selected (17.822 seconds)

```

```
select max(metro_north) from mta;
```

```

+-- selected (19.822 seconds)
0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select max(metro_north) from mta;
INFO : Compiling command(queryId=hive_20220426175421_3ce7ad43-42bb-4b3e-a203-b32a4675f2ff): se
lect max(metro_north) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:int, comment:nul
l)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426175421_3ce7ad43-42bb-4b3e-a203-b32a467
5f2ff); Time taken: 0.124 seconds
INFO : Executing command(queryId=hive_20220426175421_3ce7ad43-42bb-4b3e-a203-b32a4675f2ff): se
lect max(metro_north) from mta
WARN :
INFO : Query ID = hive_20220426175421_3ce7ad43-42bb-4b3e-a203-b32a4675f2ff
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33724
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33724/
INFO : Starting Job = job_1648648882306_33724, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33724/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33724
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 17:54:26,902 Stage-1 map = 0%, reduce = 0%
INFO : 2022-04-26 17:54:32,976 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.92 sec
INFO : 2022-04-26 17:54:40,057 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 5.24 sec
INFO : MapReduce Total cumulative CPU time: 5 seconds 240 msec
INFO : Ended Job = job_1648648882306_33724
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 5.24 sec HDFS Read: 72827 HDFS Wri
te: 106 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 5 seconds 240 msec
INFO : Completed executing command(queryId=hive_20220426175421_3ce7ad43-42bb-4b3e-a203-b32a467
5f2ff); Time taken: 19.688 seconds
INFO : OK
+-----+
| _c0 |
+-----+
| 150800 |
+-----+
1 row selected (19.824 seconds)
0: jdbc:hive2://hm-1.hpc.nyu.edu:10000>
select avg(metro_north) from mta;
```

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select avg(metro_north) from mta;
INFO  : Compiling command(queryId=hive_20220426175517_1ca7cf3e-8e22-4a15-b9c1-044a5a94824a): se
lect avg(metro_north) from mta
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20220426175517_1ca7cf3e-8e22-4a15-b9c1-044a5a9
4824a); Time taken: 0.12 seconds
INFO  : Executing command(queryId=hive_20220426175517_1ca7cf3e-8e22-4a15-b9c1-044a5a94824a): se
lect avg(metro_north) from mta
WARN  :
INFO  : Query ID = hive_20220426175517_1ca7cf3e-8e22-4a15-b9c1-044a5a94824a
INFO  : Total jobs = 1
INFO  : Launching Job 1 out of 1
INFO  : Starting task [Stage-1:MAPRED] in serial mode
INFO  : Number of reduce tasks determined at compile time: 1
INFO  : In order to change the average load for a reducer (in bytes):
INFO  :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO  : In order to limit the maximum number of reducers:
INFO  :   set hive.exec.reducers.max=<number>
INFO  : In order to set a constant number of reducers:
INFO  :   set mapreduce.job.reduces=<number>
INFO  : number of splits:1
INFO  : Submitting tokens for job: job_1648648882306_33728
INFO  : Executing with tokens: []
INFO  : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33728/
INFO  : Starting Job = job_1648648882306_33728, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33728/
INFO  : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33728
INFO  : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO  : 2022-04-26 17:55:22,542 Stage-1 map = 0%,  reduce = 0%
INFO  : 2022-04-26 17:55:28,619 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.87 sec
INFO  : 2022-04-26 17:55:34,696 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 4.78 sec
INFO  : MapReduce Total cumulative CPU time: 4 seconds 780 msec
INFO  : Ended Job = job_1648648882306_33728
INFO  : MapReduce Jobs Launched:
INFO  : Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 4.78 sec  HDFS Read: 72971 HDFS Wri
te: 117 HDFS EC Read: 0 SUCCESS
INFO  : Total MapReduce CPU Time Spent: 4 seconds 780 msec
INFO  : Completed executing command(queryId=hive_20220426175517_1ca7cf3e-8e22-4a15-b9c1-044a5a9
4824a); Time taken: 18.531 seconds
INFO  : OK
+-----+
|      _c0      |
+-----+
| 68036.44986449865 |
+-----+
1 row selected (18.664 seconds)
a: jdbc:hive2://hm-1.hpc.nyu.edu:10000/`
```

select variance(metro_north) from mta;

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select variance(metro_north) from mta;
INFO : Compiling command(queryId=hive_20220426175620_56140ba5-2b78-457f-85d5-27fb1a3f0cf0): se
lect variance(metro_north) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426175620_56140ba5-2b78-457f-85d5-27fb1a3
f0cf0); Time taken: 0.105 seconds
INFO : Executing command(queryId=hive_20220426175620_56140ba5-2b78-457f-85d5-27fb1a3f0cf0): se
lect variance(metro_north) from mta
WARN :
INFO : Query ID = hive_20220426175620_56140ba5-2b78-457f-85d5-27fb1a3f0cf0
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33731
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33731/
INFO : Starting Job = job_1648648882306_33731, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33731/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33731
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 17:56:26,891 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 17:56:31,966 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.91 sec
INFO : 2022-04-26 17:56:37,026 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 5.53 sec
INFO : MapReduce Total cumulative CPU time: 5 seconds 530 msec
INFO : Ended Job = job_1648648882306_33731
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 5.53 sec HDFS Read: 73089 HDFS Wri
te: 120 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 5 seconds 530 msec
INFO : Completed executing command(queryId=hive_20220426175620_56140ba5-2b78-457f-85d5-27fb1a3
f0cf0); Time taken: 17.29 seconds
INFO : OK
+-----+
|      _c0      |
+-----+
| 1.2094668692393563E9 |
+-----+
1 row selected (17.407 seconds)

```

```
select stddev_pop(metro_north) from mta;
```

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select stddev_pop(metro_north) from mta;
INFO : Compiling command(queryId=hive_20220426175724_28877954-5be6-462a-b867-9c5dfd258e3c): se
lect stddev_pop(metro_north) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426175724_28877954-5be6-462a-b867-9c5dfd2
58e3c); Time taken: 0.107 seconds
INFO : Executing command(queryId=hive_20220426175724_28877954-5be6-462a-b867-9c5dfd258e3c): se
lect stddev_pop(metro_north) from mta
WARN :
INFO : Query ID = hive_20220426175724_28877954-5be6-462a-b867-9c5dfd258e3c
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33734
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33734/
INFO : Starting Job = job_1648648882306_33734, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33734/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/hac
oop job -kill job_1648648882306_33734
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 17:57:30,928 Stage-1 map = 0%, reduce = 0%
INFO : 2022-04-26 17:57:35,997 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.93 sec
INFO : 2022-04-26 17:57:42,086 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 5.66 sec
INFO : MapReduce Total cumulative CPU time: 5 seconds 660 msec
INFO : Ended Job = job_1648648882306_33734
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 5.66 sec HDFS Read: 73093 HDFS Wri
te: 117 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 5 seconds 660 msec
INFO : Completed executing command(queryId=hive_20220426175724_28877954-5be6-462a-b867-9c5dfd2
58e3c); Time taken: 18.436 seconds
INFO : OK
+-----+
|      _c0      |
+-----+
| 34777.39020167207 |
+-----+
1 row selected (18.556 seconds)

select min(metro_percent) from mta;

```

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select min(metro_percent) from mta;
INFO : Compiling command(queryId=hive_20220426175757_d0029d94-6042-4418-824f-970b9c04dc4a): se
lect min(metro_percent) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426175757_d0029d94-6042-4418-824f-970b9c0
4dc4a); Time taken: 0.115 seconds
INFO : Executing command(queryId=hive_20220426175757_d0029d94-6042-4418-824f-970b9c04dc4a): se
lect min(metro_percent) from mta
WARN :
INFO : Query ID = hive_20220426175757_d0029d94-6042-4418-824f-970b9c04dc4a
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33735
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33735/
INFO : Starting Job = job_1648648882306_33735, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33735/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33735
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 17:58:03,780 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 17:58:08,910 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 2.01 sec
INFO : 2022-04-26 17:58:14,985 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 7.37 sec
INFO : MapReduce Total cumulative CPU time: 7 seconds 370 msec
INFO : Ended Job = job_1648648882306_33735
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 7.37 sec  HDFS Read: 72729 HDFS Wri
te: 104 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 7 seconds 370 msec
INFO : Completed executing command(queryId=hive_20220426175757_d0029d94-6042-4418-824f-970b9c0
4dc4a); Time taken: 18.44 seconds
INFO : OK
+---+
| _c0 |
+---+
| 0.03 |
+---+
1 row selected (18.569 seconds)

```

select max(metro_percent) from mta;

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select max(metro_percent) from mta;
INFO : Compiling command(queryId=hive_20220426175852_957631ff-aea4-41ea-a6a7-bb0ff9574760): se
lect max(metro_percent) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426175852_957631ff-aea4-41ea-a6a7-bb0ff95
74760); Time taken: 0.103 seconds
INFO : Executing command(queryId=hive_20220426175852_957631ff-aea4-41ea-a6a7-bb0ff9574760): se
lect max(metro_percent) from mta
WARN :
INFO : Query ID = hive_20220426175852_957631ff-aea4-41ea-a6a7-bb0ff9574760
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33737
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33737/
INFO : Starting Job = job_1648648882306_33737, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33737/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33737
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 17:58:58,840 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 17:59:03,906 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.92 sec
INFO : 2022-04-26 17:59:09,998 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 4.75 sec
INFO : MapReduce Total cumulative CPU time: 4 seconds 750 msec
INFO : Ended Job = job_1648648882306_33737
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 4.75 sec  HDFS Read: 72833 HDFS Wri
te: 103 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 4 seconds 750 msec
INFO : Completed executing command(queryId=hive_20220426175852_957631ff-aea4-41ea-a6a7-bb0ff95
74760); Time taken: 19.264 seconds
INFO : OK
+---+
| _c0 |
+---+
| 0.9 |
+---+
1 row selected (19.378 seconds)

```

select avg(metro_percent) from mta;

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select avg(metro_percent) from mta;
INFO : Compiling command(queryId=hive_20220426175931_b6053c59-4f07-4812-bccf-42a678c0dfc5): se
lect avg(metro_percent) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426175931_b6053c59-4f07-4812-bccf-42a678c
0dfc5); Time taken: 0.114 seconds
INFO : Executing command(queryId=hive_20220426175931_b6053c59-4f07-4812-bccf-42a678c0dfc5): se
lect avg(metro_percent) from mta
WARN :
INFO : Query ID = hive_20220426175931_b6053c59-4f07-4812-bccf-42a678c0dfc5
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33739
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_1648648882306_33739/
INFO : Starting Job = job_1648648882306_33739, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33739/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/hac
oop job -kill job_1648648882306_33739
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 17:59:36,142 Stage-1 map = 0%, reduce = 0%
INFO : 2022-04-26 17:59:41,212 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.91 sec
INFO : 2022-04-26 17:59:47,288 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6.09 sec
INFO : MapReduce Total cumulative CPU time: 6 seconds 90 msec
INFO : Ended Job = job_1648648882306_33739
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.09 sec HDFS Read: 73099 HDFS Wr
ite: 119 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 90 msec
INFO : Completed executing command(queryId=hive_20220426175931_b6053c59-4f07-4812-bccf-42a678c
0dfc5); Time taken: 17.224 seconds
INFO : OK
+-----+
|      _c0      |
+-----+
| 0.33395663956639565 |
+-----+
1 row selected (17.348 seconds)

select variance(metro_percent) from mta;

```

```

1 row selected (58.852 seconds)
0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select variance(metro_percent) from mta;
INFO  : Compiling command(queryId=hive_20220426180019_490bdebc-ebc0-4756-b773-68e4515b55bb): se
lect variance(metro_percent) from mta
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20220426180019_490bdebc-ebc0-4756-b773-68e4515
b55bb); Time taken: 0.117 seconds
INFO  : Executing command(queryId=hive_20220426180019_490bdebc-ebc0-4756-b773-68e4515b55bb): se
lect variance(metro_percent) from mta
WARN  :
INFO  : Query ID = hive_20220426180019_490bdebc-ebc0-4756-b773-68e4515b55bb
INFO  : Total jobs = 1
INFO  : Launching Job 1 out of 1
INFO  : Starting task [Stage-1:MAPRED] in serial mode
INFO  : Number of reduce tasks determined at compile time: 1
INFO  : In order to change the average load for a reducer (in bytes):
INFO  :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO  : In order to limit the maximum number of reducers:
INFO  :   set hive.exec.reducers.max=<number>
INFO  : In order to set a constant number of reducers:
INFO  :   set mapreduce.job.reduces=<number>
INFO  : number of splits:1
INFO  : Submitting tokens for job: job_1648648882306_33741
INFO  : Executing with tokens: []
INFO  : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33741/
INFO  : Starting Job = job_1648648882306_33741, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33741/
INFO  : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33741
INFO  : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO  : 2022-04-26 18:00:25,405 Stage-1 map = 0%,  reduce = 0%
INFO  : 2022-04-26 18:00:30,477 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.9 sec
INFO  : 2022-04-26 18:00:46,670 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 4.32 sec
INFO  : MapReduce Total cumulative CPU time: 5 seconds 380 msec
INFO  : Ended Job = job_1648648882306_33741
INFO  : MapReduce Jobs Launched:
INFO  : Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 5.38 sec  HDFS Read: 73105 HDFS Wri
te: 119 HDFS EC Read: 0 SUCCESS
INFO  : Total MapReduce CPU Time Spent: 5 seconds 380 msec
INFO  : Completed executing command(queryId=hive_20220426180019_490bdebc-ebc0-4756-b773-68e4515
b55bb); Time taken: 58.722 seconds
INFO  : OK
+-----+
|      _c0      |
+-----+
| 0.03564532061309763 |
+-----+
1 row selected (58.852 seconds)

```

select stddev_pop(metro_percent) from mta;

```

+----+
|   _c0   |
+----+
| 0.18879968382679466 |
+----+
1 row selected (18.325 seconds)

}: jdbc:hive2://hm-1.hpc.nyu.edu:10000/> select stddev_pop(metro_percent) from mta;
[INFO : Compiling command(queryId=hive_20220426180138_28db1a1f-f70c-44fa-8962-2bf9f04d6b9f): se
lect stddev_pop(metro_percent) from mta
[INFO : Semantic Analysis Completed
[INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
[INFO : Completed compiling command(queryId=hive_20220426180138_28db1a1f-f70c-44fa-8962-2bf9f04
d6b9f); Time taken: 0.099 seconds
[INFO : Executing command(queryId=hive_20220426180138_28db1a1f-f70c-44fa-8962-2bf9f04d6b9f): se
lect stddev_pop(metro_percent) from mta
[VARN :
[INFO : Query ID = hive_20220426180138_28db1a1f-f70c-44fa-8962-2bf9f04d6b9f
[INFO : Total jobs = 1
[INFO : Launching Job 1 out of 1
[INFO : Starting task [Stage-1:MAPRED] in serial mode
[INFO : Number of reduce tasks determined at compile time: 1
[INFO : In order to change the average load for a reducer (in bytes):
[INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
[INFO : In order to limit the maximum number of reducers:
[INFO :   set hive.exec.reducers.max=<number>
[INFO : In order to set a constant number of reducers:
[INFO :   set mapreduce.job.reduces=<number>
[INFO : number of splits:1
[INFO : Submitting tokens for job: job_1648648882306_33744
[INFO : Executing with tokens: []
[INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33744/
[INFO : Starting Job = job_1648648882306_33744, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33744/
[INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33744
[INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
[INFO : 2022-04-26 18:01:44,295 Stage-1 map = 0%,  reduce = 0%
[INFO : 2022-04-26 18:01:49,361 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.94 sec
[INFO : 2022-04-26 18:01:55,441 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 5.44 sec
[INFO : MapReduce Total cumulative CPU time: 5 seconds 440 msec
[INFO : Ended Job = job_1648648882306_33744
[INFO : MapReduce Jobs Launched:
[INFO : Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 5.44 sec  HDFS Read: 73111 HDFS Wri
te: 119 HDFS EC Read: 0 SUCCESS
[INFO : Total MapReduce CPU Time Spent: 5 seconds 440 msec
[INFO : Completed executing command(queryId=hive_20220426180138_28db1a1f-f70c-44fa-8962-2bf9f04
d6b9f); Time taken: 18.214 seconds
[INFO : OK
+----+
|   _c0   |
+----+
| 0.18879968382679466 |
+----+
1 row selected (18.325 seconds)

```

select min(access) from mta;

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000/> select min(access) from mta;
INFO : Compiling command(queryId=hive_20220426180210_e3f8bd0b-8925-4a9c-8455-d3059863bb2e): se
lect min(access) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:int, comment:nul
l)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426180210_e3f8bd0b-8925-4a9c-8455-d305986
3bb2e); Time taken: 0.113 seconds
INFO : Executing command(queryId=hive_20220426180210_e3f8bd0b-8925-4a9c-8455-d3059863bb2e): se
lect min(access) from mta
WARN :
INFO : Query ID = hive_20220426180210_e3f8bd0b-8925-4a9c-8455-d3059863bb2e
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33745
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33745/
INFO : Starting Job = job_1648648882306_33745, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33745/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33745
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 18:02:15,916 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 18:02:21,992 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 2.12 sec
INFO : 2022-04-26 18:02:27,051 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 4.97 sec
INFO : MapReduce Total cumulative CPU time: 4 seconds 970 msec
INFO : Ended Job = job_1648648882306_33745
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 4.97 sec  HDFS Read: 72818 HDFS Wri
te: 104 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 4 seconds 970 msec
INFO : Completed executing command(queryId=hive_20220426180210_e3f8bd0b-8925-4a9c-8455-d305986
3bb2e); Time taken: 17.224 seconds
INFO : OK
+---+
| _c0 |
+---+
| 2506 |
+---+
1 row selected (17.347 seconds)

```

select max(access) from mta;

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000/> select max(access) from mta;
INFO : Compiling command(queryId=hive_20220426180246_72aec4bffea1-439b-9f67-47a8632e7d7d): se
lect max(access) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:int, comment:nul
l)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426180246_72aec4bffea1-439b-9f67-47a8632
e7d7d); Time taken: 0.139 seconds
INFO : Executing command(queryId=hive_20220426180246_72aec4bffea1-439b-9f67-47a8632e7d7d): se
lect max(access) from mta
WARN :
INFO : Query ID = hive_20220426180246_72aec4bffea1-439b-9f67-47a8632e7d7d
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33748
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33748/
INFO : Starting Job = job_1648648882306_33748, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33748/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33748
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 18:02:52,676 Stage-1 map = 0%, reduce = 0%
INFO : 2022-04-26 18:02:57,743 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.91 sec
INFO : 2022-04-26 18:03:03,816 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 4.69 sec
INFO : MapReduce Total cumulative CPU time: 4 seconds 690 msec
INFO : Ended Job = job_1648648882306_33748
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.69 sec HDFS Read: 72714 HDFS Wri
te: 105 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 4 seconds 690 msec
INFO : Completed executing command(queryId=hive_20220426180246_72aec4bffea1-439b-9f67-47a8632
e7d7d); Time taken: 18.221 seconds
INFO : OK
+-----+
| _c0 |
+-----+
| 25847 |
+-----+
1 row selected (18.372 seconds)

select avg(access) from mta;

```

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000/> select avg(access) from mta;      ]
INFO  : Compiling command(queryId=hive_20220426180319_626bdd14-f95d-4681-9564-c65c9c7e167c): se
lect avg(access) from mta
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20220426180319_626bdd14-f95d-4681-9564-c65c9c7
e167c); Time taken: 0.088 seconds
INFO  : Executing command(queryId=hive_20220426180319_626bdd14-f95d-4681-9564-c65c9c7e167c): se
lect avg(access) from mta
WARN  :
INFO  : Query ID = hive_20220426180319_626bdd14-f95d-4681-9564-c65c9c7e167c
INFO  : Total jobs = 1
INFO  : Launching Job 1 out of 1
INFO  : Starting task [Stage-1:MAPRED] in serial mode
INFO  : Number of reduce tasks determined at compile time: 1
INFO  : In order to change the average load for a reducer (in bytes):
INFO  :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO  : In order to limit the maximum number of reducers:
INFO  :   set hive.exec.reducers.max=<number>
INFO  : In order to set a constant number of reducers:
INFO  :   set mapreduce.job.reduces=<number>
INFO  : number of splits:1
INFO  : Submitting tokens for job: job_1648648882306_33749
INFO  : Executing with tokens: []
INFO  : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33749/
INFO  : Starting Job = job_1648648882306_33749, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33749/
INFO  : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33749
INFO  : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO  : 2022-04-26 18:03:24,440 Stage-1 map = 0%,  reduce = 0%
INFO  : 2022-04-26 18:03:30,519 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.86 sec
INFO  : 2022-04-26 18:03:36,589 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 4.84 sec
INFO  : MapReduce Total cumulative CPU time: 4 seconds 840 msec
INFO  : Ended Job = job_1648648882306_33749
INFO  : MapReduce Jobs Launched:
INFO  : Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 4.84 sec  HDFS Read: 72962 HDFS Wri
te: 116 HDFS EC Read: 0 SUCCESS
INFO  : Total MapReduce CPU Time Spent: 4 seconds 840 msec
INFO  : Completed executing command(queryId=hive_20220426180319_626bdd14-f95d-4681-9564-c65c9c7
e167c); Time taken: 18.34 seconds
INFO  : OK
+-----+
|      _c0      |
+-----+
| 16986.9945799458 |
+-----+
1 row selected (18.439 seconds)

select variance(access) from mta;
```

```

+--+
|   _c0    |
+--+
| 3.5271354249293126E7 |
+--+
1 row selected (17.305 seconds)

```

select stddev_pop(access) from mta;

```

[0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select variance(access) from mta;
INFO : Compiling command(queryId=hive_20220426180353_659f3076-0f62-4426-b25b-b6b8645d438e): se
lect variance(access) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426180353_659f3076-0f62-4426-b25b-b6b8645
d438e); Time taken: 0.093 seconds
INFO : Executing command(queryId=hive_20220426180353_659f3076-0f62-4426-b25b-b6b8645d438e): se
lect variance(access) from mta
WARN :
INFO : Query ID = hive_20220426180353_659f3076-0f62-4426-b25b-b6b8645d438e
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33751
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33751/
INFO : Starting Job = job_1648648882306_33751, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33751/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33751
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 18:03:59,197 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 18:04:04,260 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.86 sec
INFO : 2022-04-26 18:04:09,317 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 6.38 sec
INFO : MapReduce Total cumulative CPU time: 6 seconds 380 msec
INFO : Ended Job = job_1648648882306_33751
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 6.38 sec  HDFS Read: 73078 HDFS Wri
te: 120 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 380 msec
INFO : Completed executing command(queryId=hive_20220426180353_659f3076-0f62-4426-b25b-b6b8645
d438e); Time taken: 17.199 seconds
INFO : OK
+-----+
|       _c0      |
+-----+
| 3.5271354249293126E7 |
+-----+
1 row selected (17.305 seconds)

```

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select stddev_pop(access) from mta;
INFO : Compiling command(queryId=hive_20220426180436_a2c22247-a60c-43d2-aea5-86a9e9e96b5a): se
lect stddev_pop(access) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426180436_a2c22247-a60c-43d2-aea5-86a9e9
96b5a); Time taken: 0.093 seconds
INFO : Executing command(queryId=hive_20220426180436_a2c22247-a60c-43d2-aea5-86a9e9e96b5a): se
lect stddev_pop(access) from mta
WARN :
INFO : Query ID = hive_20220426180436_a2c22247-a60c-43d2-aea5-86a9e9e96b5a
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33754
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33754/
INFO : Starting Job = job_1648648882306_33754, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33754/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33754
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 18:04:42,589 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 18:04:47,658 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.94 sec
INFO : 2022-04-26 18:04:53,728 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 5.52 sec
INFO : MapReduce Total cumulative CPU time: 5 seconds 520 msec
INFO : Ended Job = job_1648648882306_33754
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 5.52 sec HDFS Read: 73084 HDFS Wri
te: 117 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 5 seconds 520 msec
INFO : Completed executing command(queryId=hive_20220426180436_a2c22247-a60c-43d2-aea5-86a9e9
96b5a); Time taken: 18.372 seconds
INFO : OK
+-----+
|      _c0      |
+-----+
| 5938.969123450056 |
+-----+
1 row selected (18.476 seconds)

```

select min(access_percent) from mta;

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000/> select min(access_percent) from mta;
INFO : Compiling command(queryId=hive_20220426180522_74bf9cb2-97c1-4b81-8c31-e71804d342a0): se
lect min(access_percent) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:float, comment:n
ull)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426180522_74bf9cb2-97c1-4b81-8c31-e71804d
342a0); Time taken: 0.076 seconds
INFO : Executing command(queryId=hive_20220426180522_74bf9cb2-97c1-4b81-8c31-e71804d342a0): se
lect min(access_percent) from mta
WARN :
INFO : Query ID = hive_20220426180522_74bf9cb2-97c1-4b81-8c31-e71804d342a0
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33755
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33755/
INFO : Starting Job = job_1648648882306_33755, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33755/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33755
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 18:05:28,884 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 18:05:33,950 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.93 sec
INFO : 2022-04-26 18:05:40,031 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 5.85 sec
INFO : MapReduce Total cumulative CPU time: 5 seconds 850 msec
INFO : Ended Job = job_1648648882306_33755
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 5.85 sec HDFS Read: 72748 HDFS Wri
te: 105 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 5 seconds 850 msec
INFO : Completed executing command(queryId=hive_20220426180522_74bf9cb2-97c1-4b81-8c31-e71804d
342a0); Time taken: 18.116 seconds
INFO : OK
+-----+
| _c0 |
+-----+
| 0.136 |
+-----+
1 row selected (18.336 seconds)
2. idh@hive2:~/hm-1.hpc.nyu.edu..10000/>
select max(access_percent) from mta;
```

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select max(access_percent) from mta;
INFO : Compiling command(queryId=hive_20220426180558_0be88535-dfdf-4f4b-a007-159217b5e2b4): se
lect max(access_percent) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:float, comment:n
ull)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426180558_0be88535-dfdf-4f4b-a007-159217b
5e2b4); Time taken: 0.077 seconds
INFO : Executing command(queryId=hive_20220426180558_0be88535-dfdf-4f4b-a007-159217b5e2b4): se
lect max(access_percent) from mta
WARN :
INFO : Query ID = hive_20220426180558_0be88535-dfdf-4f4b-a007-159217b5e2b4
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33759
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33759/
INFO : Starting Job = job_1648648882306_33759, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33759/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33759
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 18:06:04,499 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 18:06:09,576 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.93 sec
INFO : 2022-04-26 18:06:15,648 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 6.66 sec
INFO : MapReduce Total cumulative CPU time: 6 seconds 660 msec
INFO : Ended Job = job_1648648882306_33759
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 6.66 sec  HDFS Read: 72852 HDFS Wri
te: 105 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 660 msec
INFO : Completed executing command(queryId=hive_20220426180558_0be88535-dfdf-4f4b-a007-159217b
5e2b4); Time taken: 18.272 seconds
INFO : OK
+-----+
| _c0 |
+-----+
| 1.055 |
+-----+
1 row selected (18.361 seconds)

select avg(access_percent) from mta;

```

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select avg(access_percent) from mta;
INFO : Compiling command(queryId=hive_20220426180639_4e1ff535-aac4-41dd-bcf9-1a514defd2b0): se
lect avg(access_percent) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426180639_4e1ff535-aac4-41dd-bcf9-1a514de
fd2b0); Time taken: 0.084 seconds
INFO : Executing command(queryId=hive_20220426180639_4e1ff535-aac4-41dd-bcf9-1a514defd2b0): se
lect avg(access_percent) from mta
WARN :
INFO : Query ID = hive_20220426180639_4e1ff535-aac4-41dd-bcf9-1a514defd2b0
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33761
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33761/
INFO : Starting Job = job_1648648882306_33761, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33761/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33761
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 18:06:45,178 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 18:06:50,248 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.92 sec
INFO : 2022-04-26 18:06:56,323 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 6.66 sec
INFO : MapReduce Total cumulative CPU time: 6 seconds 660 msec
INFO : Ended Job = job_1648648882306_33761
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 6.66 sec  HDFS Read: 73102 HDFS Wri
te: 118 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 660 msec
INFO : Completed executing command(queryId=hive_20220426180639_4e1ff535-aac4-41dd-bcf9-1a514de
fd2b0); Time taken: 18.149 seconds
INFO : OK
+-----+
|      _c0      |
+-----+
| 0.6514756099843397 |
+-----+
1 row selected (18.243 seconds)

select variance(access_percent) from mta;
```

```

[0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select variance(access_percent) from mta;      ]
INFO  : Compiling command(queryId=hive_20220426180722_ff729bf8-4b80-4f91-af66-15e3541faf92): se
lect variance(access_percent) from mta
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20220426180722_ff729bf8-4b80-4f91-af66-15e3541
faf92); Time taken: 0.099 seconds
INFO  : Executing command(queryId=hive_20220426180722_ff729bf8-4b80-4f91-af66-15e3541faf92): se
lect variance(access_percent) from mta
WARN  :
INFO  : Query ID = hive_20220426180722_ff729bf8-4b80-4f91-af66-15e3541faf92
INFO  : Total jobs = 1
INFO  : Launching Job 1 out of 1
INFO  : Starting task [Stage-1:MAPRED] in serial mode
INFO  : Number of reduce tasks determined at compile time: 1
INFO  : In order to change the average load for a reducer (in bytes):
INFO  :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO  : In order to limit the maximum number of reducers:
INFO  :   set hive.exec.reducers.max=<number>
INFO  : In order to set a constant number of reducers:
INFO  :   set mapreduce.job.reduces=<number>
INFO  : number of splits:1
INFO  : Submitting tokens for job: job_1648648882306_33762
INFO  : Executing with tokens: []
INFO  : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33762/
INFO  : Starting Job = job_1648648882306_33762, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33762/
INFO  : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33762
INFO  : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO  : 2022-04-26 18:07:27,445 Stage-1 map = 0%,  reduce = 0%
INFO  : 2022-04-26 18:07:33,521 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.95 sec
INFO  : 2022-04-26 18:07:39,591 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 5.92 sec
INFO  : MapReduce Total cumulative CPU time: 5 seconds 920 msec
INFO  : Ended Job = job_1648648882306_33762
INFO  : MapReduce Jobs Launched:
INFO  : Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 5.92 sec  HDFS Read: 73106 HDFS Wri
te: 120 HDFS EC Read: 0 SUCCESS
INFO  : Total MapReduce CPU Time Spent: 5 seconds 920 msec
INFO  : Completed executing command(queryId=hive_20220426180722_ff729bf8-4b80-4f91-af66-15e3541
faf92); Time taken: 18.349 seconds
INFO  : OK
+-----+
|       _c0      |
+-----+
| 0.025752989198315455 |
+-----+
1 row selected (18.459 seconds)

select stddev_pop(access_percent) from mta;
```

```

+ 1 row selected (46.874 seconds)
0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select stddev_pop(access_percent) from mta;
INFO : Compiling command(queryId=hive_20220426180833_b1962da8-3349-4633-8ad7-185aa21a98fe): se
lect stddev_pop(access_percent) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426180833_b1962da8-3349-4633-8ad7-185aa21
a98fe); Time taken: 26.893 seconds
INFO : Executing command(queryId=hive_20220426180833_b1962da8-3349-4633-8ad7-185aa21a98fe): se
lect stddev_pop(access_percent) from mta
WARN :
INFO : Query ID = hive_20220426180833_b1962da8-3349-4633-8ad7-185aa21a98fe
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33767
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33767/
INFO : Starting Job = job_1648648882306_33767, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33767/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33767
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 18:09:06,793 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 18:09:11,861 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.99 sec
INFO : 2022-04-26 18:09:17,934 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 4.95 sec
INFO : MapReduce Total cumulative CPU time: 4 seconds 950 msec
INFO : Ended Job = job_1648648882306_33767
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 4.95 sec   HDFS Read: 73006 HDFS Wri
te: 119 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 4 seconds 950 msec
INFO : Completed executing command(queryId=hive_20220426180833_b1962da8-3349-4633-8ad7-185aa21
a98fe); Time taken: 19.97 seconds
INFO : OK
+-----+
|      _c0      |
+-----+
| 0.16047737908601153 |
+-----+
1 row selected (46.874 seconds)

```

select min(traffic) from mta;

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select min(traffic) from mta;
INFO : Compiling command(queryId=hive_20220426180938_b2ecb8c5-78d8-4586-9ceb-e634fad65914): s
elect min(traffic) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:int, comment:n
ull)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426180938_b2ecb8c5-78d8-4586-9ceb-e634fad
65914); Time taken: 0.145 seconds
INFO : Executing command(queryId=hive_20220426180938_b2ecb8c5-78d8-4586-9ceb-e634fad65914): s
elect min(traffic) from mta
WARN :
INFO : Query ID = hive_20220426180938_b2ecb8c5-78d8-4586-9ceb-e634fad65914
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33769
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33769/
INFO : Starting Job = job_1648648882306_33769, Tracking URL = http://horton.hpc.nyu.edu:8088/|_
roxy/application_1648648882306_33769/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33769
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 18:09:45,415 Stage-1 map = 0%, reduce = 0%
INFO : 2022-04-26 18:09:50,545 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.98 sec
INFO : 2022-04-26 18:09:56,623 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 4.87 sec
INFO : MapReduce Total cumulative CPU time: 4 seconds 870 msec
INFO : Ended Job = job_1648648882306_33769
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.87 sec HDFS Read: 72820 HDFS Wr
ite: 106 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 4 seconds 870 msec
INFO : Completed executing command(queryId=hive_20220426180938_b2ecb8c5-78d8-4586-9ceb-e634fad
65914); Time taken: 19.031 seconds
INFO : OK
+-----+
| _c0 |
+-----+
| 156759 |
+-----+
1 row selected (19.187 seconds)

```

select max(traffic) from mta;

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000/> select max(traffic) from mta;
INFO : Compiling command(queryId=hive_20220426181010_254dfda7-80d4-4f4c-aa2f-5a039a917d84): se
lect max(traffic) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:int, comment:nul
l)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426181010_254dfda7-80d4-4f4c-aa2f-5a039a9
17d84); Time taken: 0.117 seconds
INFO : Executing command(queryId=hive_20220426181010_254dfda7-80d4-4f4c-aa2f-5a039a917d84): se
lect max(traffic) from mta
WARN :
INFO : Query ID = hive_20220426181010_254dfda7-80d4-4f4c-aa2f-5a039a917d84
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33770
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
5_33770/
INFO : Starting Job = job_1648648882306_33770, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33770/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33770
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 18:10:17,162 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 18:10:23,255 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.86 sec
INFO : 2022-04-26 18:10:29,341 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 5.91 sec
INFO : MapReduce Total cumulative CPU time: 5 seconds 910 msec
INFO : Ended Job = job_1648648882306_33770
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 5.91 sec  HDFS Read: 72820 HDFS Wri
te: 107 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 5 seconds 910 msec
INFO : Completed executing command(queryId=hive_20220426181010_254dfda7-80d4-4f4c-aa2f-5a039a9
17d84); Time taken: 19.548 seconds
INFO : OK
+-----+
| _c0   |
+-----+
| 1014806 |
+-----+
select avg(traffic) from mta;

```

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select avg(traffic) from mta;
INFO : Compiling command(queryId=hive_20220426181055_462f223e-7b57-44f2-8f6b-4181c2e8c0d0): se
lect avg(traffic) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426181055_462f223e-7b57-44f2-8f6b-4181c2e
8c0d0); Time taken: 0.131 seconds
INFO : Executing command(queryId=hive_20220426181055_462f223e-7b57-44f2-8f6b-4181c2e8c0d0): se
lect avg(traffic) from mta
WARN :
INFO : Query ID = hive_20220426181055_462f223e-7b57-44f2-8f6b-4181c2e8c0d0
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33771
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33771/
INFO : Starting Job = job_1648648882306_33771, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33771/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33771
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 18:11:01,886 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 18:11:06,957 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 2.11 sec
INFO : 2022-04-26 18:11:13,031 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 5.21 sec
INFO : MapReduce Total cumulative CPU time: 5 seconds 210 msec
INFO : Ended Job = job_1648648882306_33771
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 5.21 sec HDFS Read: 73067 HDFS Wri
te: 117 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 5 seconds 210 msec
INFO : Completed executing command(queryId=hive_20220426181055_462f223e-7b57-44f2-8f6b-4181c2e
8c0d0); Time taken: 19.922 seconds
INFO : OK
+-----+
| _c0 |
+-----+
| 779949.0284552845 |
+-----+
1 row selected (20.064 seconds)

```

select variance(traffic) from mta;

```

?: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select variance(traffic) from mta;
[INFO : Compiling command(queryId=hive_20220426181127_6112ef9d-435b-4470-b5f9-6d0f9aa31893): se
lect variance(traffic) from mta
[INFO : Semantic Analysis Completed
[INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
[INFO : Completed compiling command(queryId=hive_20220426181127_6112ef9d-435b-4470-b5f9-6d0f9aa
31893); Time taken: 0.109 seconds
[INFO : Executing command(queryId=hive_20220426181127_6112ef9d-435b-4470-b5f9-6d0f9aa31893): se
lect variance(traffic) from mta
[VARN :
[INFO : Query ID = hive_20220426181127_6112ef9d-435b-4470-b5f9-6d0f9aa31893
[INFO : Total jobs = 1
[INFO : Launching Job 1 out of 1
[INFO : Starting task [Stage-1:MAPRED] in serial mode
[INFO : Number of reduce tasks determined at compile time: 1
[INFO : In order to change the average load for a reducer (in bytes):
[INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
[INFO : In order to limit the maximum number of reducers:
[INFO :   set hive.exec.reducers.max=<number>
[INFO : In order to set a constant number of reducers:
[INFO :   set mapreduce.job.reduces=<number>
[INFO : number of splits:1
[INFO : Submitting tokens for job: job_1648648882306_33772
[INFO : Executing with tokens: []
[INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
5_33772/
[INFO : Starting Job = job_1648648882306_33772, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33772/
[INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33772
[INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
[INFO : 2022-04-26 18:11:34,480 Stage-1 map = 0%,  reduce = 0%
[INFO : 2022-04-26 18:11:39,549 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 2.21 sec
[INFO : 2022-04-26 18:11:44,607 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 5.79 sec
[INFO : MapReduce Total cumulative CPU time: 5 seconds 790 msec
[INFO : Ended Job = job_1648648882306_33772
[INFO : MapReduce Jobs Launched:
[INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 5.79 sec HDFS Read: 73081 HDFS Wri
te: 121 HDFS EC Read: 0 SUCCESS
[INFO : Total MapReduce CPU Time Spent: 5 seconds 790 msec
[INFO : Completed executing command(queryId=hive_20220426181127_6112ef9d-435b-4470-b5f9-6d0f9aa
31893); Time taken: 18.865 seconds
[INFO : OK
+-----+
| _c0      |
+-----+
| 2.5614269820840664E10 |
+-----+
1 row selected (18.985 seconds)

select stddev_pop(traffic) from mta;
```

```

+----+
|    _c0    |
+----+
| 160044.58697763152 |
+----+
1 row selected (19.719 seconds)

```

[0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select stddev_pop(traffic) from mta;

INFO : Compiling command(queryId=hive_20220426181207_47721538-64d9-4140-9915-378b5a3fd24d): select stddev_pop(traffic) from mta

INFO : Semantic Analysis Completed

INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:null)], properties:null)

INFO : Completed compiling command(queryId=hive_20220426181207_47721538-64d9-4140-9915-378b5a3fd24d); Time taken: 0.118 seconds

INFO : Executing command(queryId=hive_20220426181207_47721538-64d9-4140-9915-378b5a3fd24d): select stddev_pop(traffic) from mta

WARN :

INFO : Query ID = hive_20220426181207_47721538-64d9-4140-9915-378b5a3fd24d

INFO : Total jobs = 1

INFO : Launching Job 1 out of 1

INFO : Starting task [Stage-1:MAPRED] in serial mode

INFO : Number of reduce tasks determined at compile time: 1

INFO : In order to change the average load for a reducer (in bytes):

INFO : set hive.exec.reducers.bytes.per.reducer=<number>

INFO : In order to limit the maximum number of reducers:

INFO : set hive.exec.reducers.max=<number>

INFO : In order to set a constant number of reducers:

INFO : set mapreduce.job.reduces=<number>

INFO : number of splits:1

INFO : Submitting tokens for job: job_1648648882306_33773

INFO : Executing with tokens: []

INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_1648648882306_33773/

INFO : Starting Job = job_1648648882306_33773, Tracking URL = http://horton.hpc.nyu.edu:8088/proxy/application_1648648882306_33773/

INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/hadoop job -kill job_1648648882306_33773

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

INFO : 2022-04-26 18:12:13,457 Stage-1 map = 0%, reduce = 0%

INFO : 2022-04-26 18:12:20,541 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.94 sec

INFO : 2022-04-26 18:12:26,609 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 4.79 sec

INFO : MapReduce Total cumulative CPU time: 4 seconds 790 msec

INFO : Ended Job = job_1648648882306_33773

INFO : MapReduce Jobs Launched:

INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.79 sec HDFS Read: 73085 HDFS Write: 118 HDFS EC Read: 0 SUCCESS

INFO : Total MapReduce CPU Time Spent: 4 seconds 790 msec

INFO : Completed executing command(queryId=hive_20220426181207_47721538-64d9-4140-9915-378b5a3fd24d); Time taken: 19.589 seconds

INFO : OK

```
select min(traffic_percent) from mta;
```

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select min(traffic_percent) from mta;
INFO  : Compiling command(queryId=hive_20220426181239_8399151b-31bb-4181-a3e1-5011f940fc9f): se
lect min(traffic_percent) from mta
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20220426181239_8399151b-31bb-4181-a3e1-5011f94
0fc9f); Time taken: 0.116 seconds
INFO  : Executing command(queryId=hive_20220426181239_8399151b-31bb-4181-a3e1-5011f940fc9f): se
lect min(traffic_percent) from mta
WARN  :
INFO  : Query ID = hive_20220426181239_8399151b-31bb-4181-a3e1-5011f940fc9f
INFO  : Total jobs = 1
INFO  : Launching Job 1 out of 1
INFO  : Starting task [Stage-1:MAPRED] in serial mode
INFO  : Number of reduce tasks determined at compile time: 1
INFO  : In order to change the average load for a reducer (in bytes):
INFO  :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO  : In order to limit the maximum number of reducers:
INFO  :   set hive.exec.reducers.max=<number>
INFO  : In order to set a constant number of reducers:
INFO  :   set mapreduce.job.reduces=<number>
INFO  : number of splits:1
INFO  : Submitting tokens for job: job_1648648882306_33775
INFO  : Executing with tokens: []
INFO  : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33775/
INFO  : Starting Job = job_1648648882306_33775, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33775/
INFO  : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33775
INFO  : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO  : 2022-04-26 18:12:46,240 Stage-1 map = 0%,  reduce = 0%
INFO  : 2022-04-26 18:12:51,306 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 2.12 sec
INFO  : 2022-04-26 18:12:57,376 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 5.12 sec
INFO  : MapReduce Total cumulative CPU time: 5 seconds 120 msec
INFO  : Ended Job = job_1648648882306_33775
INFO  : MapReduce Jobs Launched:
INFO  : Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 5.12 sec   HDFS Read: 72837 HDFS Wri
te: 105 HDFS EC Read: 0 SUCCESS
INFO  : Total MapReduce CPU Time Spent: 5 seconds 120 msec
INFO  : Completed executing command(queryId=hive_20220426181239_8399151b-31bb-4181-a3e1-5011f94
0fc9f); Time taken: 19.355 seconds
INFO  : OK
+-----+
| _c0 |
+-----+
| 0.183 |
+-----+
1 row selected (19.484 seconds)
2. jdbc:hive2://hm-1.hpc.nyu.edu:10000/>
select max(traffic_percent) from mta;
```

```

1 row selected (19.404 seconds)
0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select max(traffic_percent) from mta;
INFO : Compiling command(queryId=hive_20220426181314_7e54c307-662e-45bb-9730-6c3b7a6695e1): se
lect max(traffic_percent) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426181314_7e54c307-662e-45bb-9730-6c3b7a6
695e1); Time taken: 0.122 seconds
INFO : Executing command(queryId=hive_20220426181314_7e54c307-662e-45bb-9730-6c3b7a6695e1): se
lect max(traffic_percent) from mta
WARN :
INFO : Query ID = hive_20220426181314_7e54c307-662e-45bb-9730-6c3b7a6695e1
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO :   set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO :   set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33776
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33776/
INFO : Starting Job = job_1648648882306_33776, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33776/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33776
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 18:13:19,962 Stage-1 map = 0%,  reduce = 0%
INFO : 2022-04-26 18:13:26,042 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.92 sec
INFO : 2022-04-26 18:13:31,110 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 6.41 sec
INFO : MapReduce Total cumulative CPU time: 6 seconds 410 msec
INFO : Ended Job = job_1648648882306_33776
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.41 sec HDFS Read: 72837 HDFS Wri
te: 105 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 410 msec
INFO : Completed executing command(queryId=hive_20220426181314_7e54c307-662e-45bb-9730-6c3b7a6
695e1); Time taken: 17.301 seconds
INFO : OK
+---+
| _c0 |
+---+
| 1.214 |
+---+
1 row selected (17.463 seconds)

```

select avg(traffic_percent) from mta;

```

-----+
): jdbc:hive2://hm-1.hpc.nyu.edu:10000/> select avg(traffic_percent) from mta;
[INFO : Compiling command(queryId=hive_20220426181345_3c19c93e-b48b-48f5-99ee-18c87a00a0f5): se
lect avg(traffic_percent) from mta
[INFO : Semantic Analysis Completed
[INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
[INFO : Completed compiling command(queryId=hive_20220426181345_3c19c93e-b48b-48f5-99ee-18c87a0
0a0f5); Time taken: 0.109 seconds
[INFO : Executing command(queryId=hive_20220426181345_3c19c93e-b48b-48f5-99ee-18c87a00a0f5): se
lect avg(traffic_percent) from mta
[VARN :
[INFO : Query ID = hive_20220426181345_3c19c93e-b48b-48f5-99ee-18c87a00a0f5
[INFO : Total jobs = 1
[INFO : Launching Job 1 out of 1
[INFO : Starting task [Stage-1:MAPRED] in serial mode
[INFO : Number of reduce tasks determined at compile time: 1
[INFO : In order to change the average load for a reducer (in bytes):
[INFO :   set hive.exec.reducers.bytes.per.reducer=<number>
[INFO : In order to limit the maximum number of reducers:
[INFO :   set hive.exec.reducers.max=<number>
[INFO : In order to set a constant number of reducers:
[INFO :   set mapreduce.job.reduces=<number>
[INFO : number of splits:1
[INFO : Submitting tokens for job: job_1648648882306_33780
[INFO : Executing with tokens: []
[INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33780/
[INFO : Starting Job = job_1648648882306_33780, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33780/
[INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33780
[INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
[INFO : 2022-04-26 18:13:50,241 Stage-1 map = 0%,  reduce = 0%
[INFO : 2022-04-26 18:13:56,319 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.89 sec
[INFO : 2022-04-26 18:14:02,393 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 5.48 sec
[INFO : MapReduce Total cumulative CPU time: 5 seconds 480 msec
[INFO : Ended Job = job_1648648882306_33780
[INFO : MapReduce Jobs Launched:
[INFO : Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 5.48 sec  HDFS Read: 73102 HDFS Wri
te: 118 HDFS EC Read: 0 SUCCESS
[INFO : Total MapReduce CPU Time Spent: 5 seconds 480 msec
[INFO : Completed executing command(queryId=hive_20220426181345_3c19c93e-b48b-48f5-99ee-18c87a0
0a0f5); Time taken: 18.322 seconds
[INFO : OK
-----+
|      _c0      |
-----+
| 0.8505460704607051  |
-----+
1 row selected (18.442 seconds)
```

```

select variance(traffic\_percent) from mta;

```

[0: jdbc:hive2://hm-1.hpc.nyu.edu:10000> select variance(traffic_percent) from mta;
INFO : Compiling command(queryId=hive_20220426181419_569bbede-c539-4ca3-bb7d-a576e305189c): se
lect variance(traffic_percent) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426181419_569bbede-c539-4ca3-bb7d-a576e30
5189c); Time taken: 0.482 seconds
INFO : Executing command(queryId=hive_20220426181419_569bbede-c539-4ca3-bb7d-a576e305189c): se
lect variance(traffic_percent) from mta
WARN :
INFO : Query ID = hive_20220426181419_569bbede-c539-4ca3-bb7d-a576e305189c
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO : set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO : set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO : set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33781
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33781/
INFO : Starting Job = job_1648648882306_33781, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33781/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33781
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 18:14:27,347 Stage-1 map = 0%, reduce = 0%
INFO : 2022-04-26 18:14:32,415 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.02 sec
INFO : 2022-04-26 18:14:38,488 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6.3 sec
INFO : MapReduce Total cumulative CPU time: 6 seconds 300 msec
INFO : Ended Job = job_1648648882306_33781
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.3 sec HDFS Read: 73110 HDFS Writ
e: 119 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 300 msec
INFO : Completed executing command(queryId=hive_20220426181419_569bbede-c539-4ca3-bb7d-a576e30
5189c); Time taken: 19.754 seconds
INFO : OK
+-----+
| _c0 |
+-----+
| 0.02814307172575116 |
+-----+
1 row selected (20.248 seconds)

select stddev_pop(traffic_percent) from mta;

```

```

0: jdbc:hive2://hm-1.hpc.nyu.edu:10000/> select stddev_pop(traffic_percent) from mta;
INFO : Compiling command(queryId=hive_20220426181457_b198cf72-677e-48c7-a06b-080545e0ff13): se
lect stddev_pop(traffic_percent) from mta
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:double, comment:
null)], properties:null)
INFO : Completed compiling command(queryId=hive_20220426181457_b198cf72-677e-48c7-a06b-080545e
0ff13); Time taken: 0.098 seconds
INFO : Executing command(queryId=hive_20220426181457_b198cf72-677e-48c7-a06b-080545e0ff13): se
lect stddev_pop(traffic_percent) from mta
WARN :
INFO : Query ID = hive_20220426181457_b198cf72-677e-48c7-a06b-080545e0ff13
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPRED] in serial mode
INFO : Number of reduce tasks determined at compile time: 1
INFO : In order to change the average load for a reducer (in bytes):
INFO : set hive.exec.reducers.bytes.per.reducer=<number>
INFO : In order to limit the maximum number of reducers:
INFO : set hive.exec.reducers.max=<number>
INFO : In order to set a constant number of reducers:
INFO : set mapreduce.job.reduces=<number>
INFO : number of splits:1
INFO : Submitting tokens for job: job_1648648882306_33783
INFO : Executing with tokens: []
INFO : The url to track the job: http://horton.hpc.nyu.edu:8088/proxy/application_164864888230
6_33783/
INFO : Starting Job = job_1648648882306_33783, Tracking URL = http://horton.hpc.nyu.edu:8088/p
roxy/application_1648648882306_33783/
INFO : Kill Command = /opt/cloudera/parcels/CDH-6.3.4-1.cdh6.3.4.p0.6626826/lib/hadoop/bin/had
oop job -kill job_1648648882306_33783
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2022-04-26 18:15:04,071 Stage-1 map = 0%, reduce = 0%
INFO : 2022-04-26 18:15:09,138 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.02 sec
INFO : 2022-04-26 18:15:15,206 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 4.78 sec
INFO : MapReduce Total cumulative CPU time: 4 seconds 780 msec
INFO : Ended Job = job_1648648882306_33783
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.78 sec HDFS Read: 73114 HDFS Wri
te: 118 HDFS EC Read: 0 SUCCESS
INFO : Total MapReduce CPU Time Spent: 4 seconds 780 msec
INFO : Completed executing command(queryId=hive_20220426181457_b198cf72-677e-48c7-a06b-080545e
0ff13); Time taken: 18.222 seconds
INFO : OK
+-----+
| _c0 |
+-----+
| 0.1677589691365298 |
+-----+
1 row selected (18.332 seconds)

```