

Code Drop 2 Screen Shots

Type of vehicle involved in accident

```
rs6785@login-1:~
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_70628
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2019-11-30 20:21:15,274 Stage-1 map = 0%, reduce = 0%
INFO : 2019-11-30 20:21:21,422 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.76 sec
INFO : 2019-11-30 20:21:28,585 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_1569350662793_70628
INFO : Launching Job 2 out of 2
INFO : Starting task [Stage->2: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_1569350662793_70630
INFO : The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_70630/
INFO : Starting Job = job_1569350662793_70630, Tracking URL = http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_70630
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_70630
INFO : Hadoop job information for Stage-2: number of mappers: 1; number of reducers: 1
INFO : 2019-11-30 20:21:43,270 Stage-2 map = 0%, reduce = 0%
INFO : 2019-11-30 20:21:48,427 Stage-2 map = 100%, reduce = 0%, Cumulative CPU 1.49 sec
INFO : 2019-11-30 20:21:54,579 Stage-2 map = 100%, reduce = 100%, Cumulative CPU 4.63 sec
INFO : MapReduce Total cumulative CPU time: 4 seconds 630 msec
INFO : Ended Job = job_1569350662793_70630
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 5.52 sec HDFS Read: 1267080 HDFS Write: 1905 SUCCESS
INFO : Stage-Stage-2: Map: 1 Reduce: 1 Cumulative CPU: 4.63 sec HDFS Read: 6844 HDFS Write: 65 SUCCESS
INFO : Total MapReduce CPU Time Spent: 10 seconds 150 msec
INFO : Completed executing command(queryId=hive_20191130202121_f6b0a509-5c00-408c-b678-b29bb38ff31d); Time taken: 53.938 seconds
INFO : OK
+-----+-----+-----+
| vehicle | total_accidents |
+-----+-----+-----+
| passenger | 12506 |
| station | 7853 |
| sport | 7853 |
| bicycle | 5074 |
| taxi | 3036 |
+-----+-----+
5 rows selected (54.393 seconds)
0: jdbc:hive2://babar.es.its.nyu.edu:10000/>
```

Station zipcode with most accidents

```
rs6785@login-1:~
INFO : 2019-12-01 20:37:56,088 Stage-4 map = 0%, reduce = 0%
INFO : 2019-12-01 20:38:02,271 Stage-4 map = 100%, reduce = 0%, Cumulative CPU 1.57 sec
INFO : 2019-12-01 20:38:09,448 Stage-4 map = 100%, reduce = 100%, Cumulative CPU 4.85 sec
INFO : MapReduce Total cumulative CPU time: 4 seconds 850 msec
INFO : Ended Job = job_1569350662793_73008
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 7.24 sec HDFS Read: 1267137 HDFS Write: 4470 SUCCESS
INFO : Stage-Stage-3: Map: 1 Reduce: 1 Cumulative CPU: 5.05 sec HDFS Read: 12405 HDFS Write: 1604 SUCCESS
INFO : Stage-Stage-4: Map: 1 Reduce: 1 Cumulative CPU: 4.85 sec HDFS Read: 6447 HDFS Write: 661 SUCCESS
INFO : Total MapReduce CPU Time Spent: 17 seconds 140 msec
INFO : Completed executing command(queryId=hive_20191201203636_eb644985-5707-4e16-8e5c-d365597ac555); Time taken: 107.087 seconds
INFO : OK
+-----+-----+
| postcode.zipcode | x.total |
+-----+-----+
| 11211 | 1404 |
| 10002 | 1095 |
| 11206 | 954 |
| 10011 | 869 |
| 11217 | 868 |
| 11237 | 826 |
| 11205 | 748 |
| 11238 | 744 |
| 10003 | 742 |
| 10009 | 669 |
| 11216 | 643 |
| 11207 | 638 |
| 11220 | 635 |
| 10001 | 630 |
| 10012 | 590 |
| 10018 | 522 |
| 1103 | 512 |
| 11201 | 509 |
| 11215 | 508 |
| 11213 | 503 |
| 11222 | 494 |
| 10026 | 480 |
| 10017 | 478 |
| 11233 | 453 |
| 11225 | 424 |
| 10023 | 417 |
| 11221 | 400 |
| 11373 | 386 |
+-----+
```

Total Accidents monthwise

```
rs0785@login-1-1:~
```

```
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 not specified. Estimated from input data size: 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_1569350662793_70601
INFO : The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_70601/
INFO : Starting Job = job_1569350662793_70601, Tracking URL = http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_70601
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_70601
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2019-11-30 20:06:54,104 Stage-1 map = 0%, reduce = 0%
INFO : 2019-11-30 20:07:00,279 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.75 sec
INFO : 2019-11-30 20:07:06,465 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_1569350662793_70601
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.04 sec HDFS Read: 1267924 HDFS Write: 87 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 40 msec
INFO : Completed executing command(queryId=hive_20191130200606_a48dc725-b492-4ac8-bb91-784e6196d2cc); Time taken: 26.787 seconds
INFO : OK
+-----+-----+
| month | total_accidents |
+-----+-----+
| 1     | 1807           |
| 2     | 1707           |
| 3     | 2150           |
| 4     | 2971           |
| 5     | 4209           |
| 6     | 4874           |
| 7     | 5679           |
| 8     | 5828           |
| 9     | 5628           |
| 10    | 4762           |
| 11    | 3272           |
| 12    | 2286           |
+-----+-----+
12 rows selected (27.205 seconds)
0: jdbc:hive2://babar.es.its.nyu.edu:10000/>
```



Yearly Fatalities

```

rs6785@login-2-1:~/Project/FilesForClean
from accidentinfo
where killed = 1
group by year
INFO : Query ID = hive_20191130235151_c22e0362-c776-4df7-bd1d-cb3af8287123
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 not specified. Estimated from input data size: 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_1569350662793_71013
INFO : The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_71013/
INFO : Starting Job = job_1569350662793_71013, Tracking URL = http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_71013/
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_71013
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2019-11-30 23:51:29,010 Stage-1 map = 0%, reduce = 0%
INFO : 2019-11-30 23:51:35,150 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.53 sec
INFO : 2019-11-30 23:51:42,315 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6.7 sec
INFO : MapReduce Total cumulative CPU time: 6 seconds 700 msec
INFO : Ended Job = job_1569350662793_71013
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.7 sec HDFS Read: 1449971 HDFS Write: 80 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 700 msec
INFO : Completed executing command(queryId=hive_20191130235151_c22e0362-c776-4df7-bd1d-cb3af8287123); Time taken: 26.069 seconds
INFO : OK
+-----+
| year | _c1 |
+-----+
| 2012 | 2382 |
| 2013 | 4393 |
| 2014 | 4380 |
| 2015 | 4693 |
| 2016 | 4737 |
| 2017 | 5682 |
| 2018 | 5387 |
| 2019 | 4892 |
+-----+
8 rows selected (26.534 seconds)
0: jdbc:hive2://babar.es.its.nyu.edu:10000/>
```

Zipcode with most accidents

```

rs6785@login-1-1:~
+-----+
| zipcode | totalkilled |
+-----+
| 11211 | 1190 |
| 10002 | 884 |
| 11206 | 792 |
| 11237 | 709 |
| 11217 | 680 |
| 11372 | 658 |
| 11205 | 630 |
| 11238 | 618 |
| 10011 | 618 |
| 11226 | 578 |
| 10003 | 546 |
| 11220 | 545 |
| 11218 | 534 |
| 11207 | 534 |
| 11216 | 514 |
| 11210 | 511 |
| 11223 | 509 |
| 10009 | 505 |
| 11219 | 488 |
| 10001 | 476 |
| 11230 | 455 |
| 11215 | 427 |
| 11103 | 426 |
| 11213 | 423 |
| 10012 | 422 |
| 11377 | 421 |
| 11368 | 408 |
| 11222 | 404 |
| 10026 | 395 |
| 11233 | 384 |
| 11201 | 381 |
| 10017 | 358 |
| 10018 | 352 |
| 11212 | 346 |
| 11225 | 344 |
| 11221 | 336 |
| 11203 | 334 |
| 11204 | 332 |
| 11229 | 326 |
| 11373 | 305 |
| 11369 | 286 |
+-----+
```

```

000, code=10004)
0: jdbc:hive2://babar.es.its.nyu.edu:10000> select COUNT (DISTINCT startid) from citibike;
INFO : Compiling command(queryId=hive_20191115154141_c1efe3f5-3fe3-41c0-b033-e01ee1ddbf87): select COUNT (DISTINCT startid) from citibike
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_20191115154141_c1efe3f5-3fe3-41c0-b033-e01ee1ddbf87); Time taken: 0.352 seconds
INFO : Executing command(queryId=hive_20191115154141_c1efe3f5-3fe3-41c0-b033-e01ee1ddbf87): select COUNT (DISTINCT startid) from citibike
INFO : Query ID = hive_20191115154141_c1efe3f5-3fe3-41c0-b033-e01ee1ddbf87
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:17
INFO : Submitting tokens for job: job_1569350662793_42307
INFO : The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_42307/
INFO : Starting Job = job_1569350662793_42307, Tracking URL = http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_42307
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_42307
INFO : Hadoop job information for Stage-1: number of mappers: 17; number of reducers: 1
INFO : 2019-11-15 15:41:48,447 Stage-1 map = 0%, reduce = 0%
INFO : 2019-11-15 15:41:54,600 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 15.52 sec
INFO : 2019-11-15 15:41:55,623 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 43.29 sec
INFO : 2019-11-15 15:41:56,645 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 78.33 sec
INFO : 2019-11-15 15:42:01,766 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 82.86 sec
INFO : MapReduce Total cumulative CPU time: 1 minutes 22 seconds 860 msec
INFO : Ended Job = job_1569350662793_42307
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 17 Reduce: 1 Cumulative CPU: 82.86 sec HDFS Read: 3675706384 HDFS Write: 4 SUCCESS
INFO : Total MapReduce CPU Time Spent: 1 minutes 22 seconds 860 msec
INFO : Completed executing command(queryId=hive_20191115154141_c1efe3f5-3fe3-41c0-b033-e01ee1ddbf87); Time taken: 29.102 seconds
INFO : OK
+---+
| _c0 |
+---+
| 951 |
+---+
1 row selected (29.605 seconds)
0: jdbc:hive2://babar.es.its.nyu.edu:10000>

```

```

YEAR(date1)
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:int, comment:null), FieldSchema(name:_c1, type:int, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20191116154242_e9beac96-8f09-43b5-aebb-24664c84be99): Time taken: 0.324 seconds
INFO : Executing command(queryId=hive_20191116154242_e9beac96-8f09-43b5-aebb-24664c84be99): select YEAR(date1),count(*) from finaldata WHERE MONTH(date1) BETWEEN 4 AND 6 group by YEAR(date1)
INFO : Query ID = hive_20191116154242_e9beac96-8f09-43b5-aebb-24664c84be99
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 not specified. Estimated from input data size: 69
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:20
INFO : Submitting tokens for job: job_1569350662793_42808
INFO : The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_42808/
INFO : Starting Job = job_1569350662793_42808, Tracking URL = http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_42808
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_42808
INFO : Hadoop job information for Stage-1: number of mappers: 20; number of reducers: 69
INFO : 2019-11-16 15:43:02,558 Stage-1 map = 0%, reduce = 0%
INFO : 2019-11-16 15:43:08,709 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 5.94 sec
INFO : 2019-11-16 15:43:09,738 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 12.26 sec
INFO : 2019-11-16 15:43:10,763 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 43.11 sec
INFO : 2019-11-16 15:43:11,787 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 175.79 sec
INFO : 2019-11-16 15:43:12,810 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 187.87 sec
INFO : 2019-11-16 15:43:16,908 Stage-1 map = 100%, reduce = 26%, Cumulative CPU 247.13 sec
INFO : 2019-11-16 15:43:18,076 Stage-1 map = 100%, reduce = 45%, Cumulative CPU 291.53 sec
INFO : 2019-11-16 15:43:19,100 Stage-1 map = 100%, reduce = 48%, Cumulative CPU 298.74 sec
INFO : 2019-11-16 15:43:20,122 Stage-1 map = 100%, reduce = 49%, Cumulative CPU 301.91 sec
INFO : 2019-11-16 15:43:21,147 Stage-1 map = 100%, reduce = 74%, Cumulative CPU 354.13 sec
INFO : 2019-11-16 15:43:22,176 Stage-1 map = 100%, reduce = 91%, Cumulative CPU 391.82 sec
INFO : 2019-11-16 15:43:23,192 Stage-1 map = 100%, reduce = 96%, Cumulative CPU 404.58 sec
INFO : 2019-11-16 15:43:24,214 Stage-1 map = 100%, reduce = 97%, Cumulative CPU 407.58 sec
INFO : 2019-11-16 15:43:25,236 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 413.28 sec
INFO : MapReduce Total cumulative CPU time: 6 minutes 53 seconds 280 msec
INFO : Ended Job = job_1569350662793_42808
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 20 Reduce: 69 Cumulative CPU: 413.28 sec HDFS Read: 4630012218 HDFS Write: 39 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 minutes 53 seconds 280 msec
INFO : Completed executing command(queryId=hive_20191116154242_e9beac96-8f09-43b5-aebb-24664c84be99); Time taken: 38.943 seconds
INFO : OK
+---+
| _c0 | _c1 |
+---+
| 2017 | 1978855 |
| 2018 | 2625746 |
| 2019 | 3018196 |
+---+
3 rows selected (39.43 seconds)

```

```

[spark@babar1]
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:int, comment:null), FieldSchema(name:_c1, type:bigint, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20191116153737_ab75f133-807a-44a1-9893-5f7638117a00); Time taken: 0.45 seconds
INFO : Executing command(queryId=hive_20191116153737_ab75f133-807a-44a1-9893-5f7638117a00): select YEAR(date1),count(*) from finaldata WHERE MONTH(date1) BETWEEN 1 AND 3 group by YEAR(date1)
INFO : Query ID = hive_20191116153737_ab75f133-807a-44a1-9893-5f7638117a00
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 not specified. Estimated from input data size: 69
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:20
INFO : Submitting tokens for job: job_1569350662793_42804
INFO : The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_42804/
INFO : Starting Job = job_1569350662793_42804, Tracking URL = http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_42804/
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_42804
INFO : Hadoop job information for Stage-1: number of mappers: 20; number of reducers: 69
INFO : 2019-11-16 15:37:14,801 Stage-1 map = 0%, reduce = 0%
INFO : 2019-11-16 15:37:20,955 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 8.46 sec
INFO : 2019-11-16 15:37:21,988 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 14.79 sec
INFO : 2019-11-16 15:37:23,003 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 49.67 sec
INFO : 2019-11-16 15:37:24,033 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 157.86 sec
INFO : 2019-11-16 15:37:25,058 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 168.57 sec
INFO : 2019-11-16 15:37:30,175 Stage-1 map = 100%, reduce = 29%, Cumulative CPU 231.21 sec
INFO : 2019-11-16 15:37:31,200 Stage-1 map = 100%, reduce = 48%, Cumulative CPU 272.36 sec
INFO : 2019-11-16 15:37:33,247 Stage-1 map = 100%, reduce = 77%, Cumulative CPU 336.29 sec
INFO : 2019-11-16 15:37:35,294 Stage-1 map = 100%, reduce = 94%, Cumulative CPU 376.64 sec
INFO : 2019-11-16 15:37:36,317 Stage-1 map = 100%, reduce = 97%, Cumulative CPU 387.21 sec
INFO : 2019-11-16 15:37:37,339 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 393.44 sec
INFO : MapReduce Total cumulative CPU time: 6 minutes 33 seconds 440 msec
INFO : Ended Job = job_1569350662793_42804
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 20 Reduce: 69 Cumulative CPU: 393.44 sec HDFS Read: 4630012218 HDFS Write: 38 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 minutes 33 seconds 440 msec
INFO : Completed executing command(queryId=hive_20191116153737_ab75f133-807a-44a1-9893-5f7638117a00); Time taken: 37.785 seconds
INFO : OK
+-----+
| _c0 | _c1 |
+-----+
| 2017 | 889024 |
| 2018 | 1053216 |
| 2019 | 1348218 |
+-----+
3 rows selected (38.384 seconds)
0: jdbc:hive2://babar.es.its.nyu.edu:10000/> ■

```

```

5 rows selected (41.272 seconds)
0: jdbc:hive2://babar.es.its.nyu.edu:10000/> select YEAR(date1), COUNT(*) from finaldata group by YEAR(date1);
INFO : Compiling command(queryId=hive_20191116145555_889ba5f0-7b51-4365-9ca9-c908db9b6fb4); select YEAR(date1), COUNT(*) from finaldata group by YEAR(date1)
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:int, comment:null), FieldSchema(name:_c1, type:bigint, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20191116145555_889ba5f0-7b51-4365-9ca9-c908db9b6fb4); Time taken: 0.357 seconds
INFO : Executing command(queryId=hive_20191116145555_889ba5f0-7b51-4365-9ca9-c908db9b6fb4); select YEAR(date1), COUNT(*) from finaldata group by YEAR(date1)
INFO : Query ID = hive_20191116145555_889ba5f0-7b51-4365-9ca9-c908db9b6fb4
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 not specified. Estimated from input data size: 69
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:20
INFO : Submitting tokens for job: job_1569350662793_42764
INFO : The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_42764/
INFO : Starting Job = job_1569350662793_42764, Tracking URL = http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_42764/
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_42764
INFO : Hadoop job information for Stage-1: number of mappers: 20; number of reducers: 69
INFO : 2019-11-16 14:55:17,008 Stage-1 map = 0%, reduce = 0%
INFO : 2019-11-16 14:55:24,843 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 22.87 sec
INFO : 2019-11-16 14:55:25,864 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 93.94 sec
INFO : 2019-11-16 14:55:26,885 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 124.18 sec
INFO : 2019-11-16 14:55:27,999 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 132.98 sec
INFO : 2019-11-16 14:55:30,991 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 140.44 sec
INFO : 2019-11-16 14:55:32,013 Stage-1 map = 100%, reduce = 20%, Cumulative CPU 200.37 sec
INFO : 2019-11-16 14:55:34,056 Stage-1 map = 100%, reduce = 45%, Cumulative CPU 237.35 sec
INFO : 2019-11-16 14:55:35,078 Stage-1 map = 100%, reduce = 48%, Cumulative CPU 244.27 sec
INFO : 2019-11-16 14:55:36,099 Stage-1 map = 100%, reduce = 77%, Cumulative CPU 309.11 sec
INFO : 2019-11-16 14:55:38,141 Stage-1 map = 100%, reduce = 93%, Cumulative CPU 345.31 sec
INFO : 2019-11-16 14:55:39,162 Stage-1 map = 100%, reduce = 97%, Cumulative CPU 355.73 sec
INFO : 2019-11-16 14:55:40,182 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 361.39 sec
INFO : MapReduce Total cumulative CPU time: 6 minutes 1 seconds 390 msec
INFO : Ended Job = job_1569350662793_42764
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 20 Reduce: 69 Cumulative CPU: 361.39 sec HDFS Read: 4629997751 HDFS Write: 62 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 minutes 1 seconds 390 msec
INFO : Completed executing command(queryId=hive_20191116145555_889ba5f0-7b51-4365-9ca9-c908db9b6fb4); Time taken: 39.802 seconds
INFO : OK
+-----+
| _c0 | _c1 |
+-----+
| NULL | 704355 |
| 2016 | 1512481 |
| 2017 | 6954513 |
| 2018 | 8573274 |
| 2019 | 8077645 |
+-----+
5 rows selected (40.271 seconds)
0: jdbc:hive2://babar.es.its.nyu.edu:10000/> ■

```

areazipcode	totalaccidents	totalstations	totaltrips
11211	42120	30	386485
10002	26280	24	1046098
10009	21408	32	1298258
11206	20988	22	113445
11217	19964	23	235992
11238	18600	25	289167
10011	17380	20	1158090
11215	14732	29	309464
11201	14252	28	485190
10003	14098	19	988688
11222	13832	28	240994
10001	11340	18	858031
11237	10738	13	56496
11231	10280	40	398846
10019	9960	30	1461734
11205	8976	12	165131
11221	8400	21	224021
11216	8359	13	106297
10029	7144	19	191730
10017	6692	14	747498
11101	6272	32	226525
10022	5742	18	597315
11207	5742	9	113665
11225	5512	13	54757
10023	5421	13	687025
10012	5310	9	528001
11106	4959	19	59676
10028	4953	13	242603
10018	4698	9	413472
10021	4511	13	362006
11102	4464	16	49090
10035	4386	17	213128
10026	3360	7	130536
10027	3185	13	98435
10025	3059	19	492391
10014	2754	18	963819
10016	2665	13	640069
11213	2515	5	101805
10013	2119	13	544069
10037	2030	10	533589
10065	1980	6	185239
10038	1568	14	419346
11103	1536	3	10589
11233	1359	3	36520
11220	1270	2	632
11105	987	7	31011
10007	924	7	335448
10199	780	3	118532
11369	664	2	10346
11208	654	2	41574
10005	584	8	407613
11109	395	5	58693

Weather Data Screenshots

```
0: jdbc:hive2://babar.es.its.nyu.edu:10000> select count(*) as no_of_days from weather_final_2 where dew < 1;
INFO : Compiling command(queryId=hive_20191201155959_adcb1ad4-624e-4d85-a15c-711355caeac8): select count(*) as no_of_days from weather_final_2 where dew < 1
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:no_of_days, type:bigint, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20191201155959_adcb1ad4-624e-4d85-a15c-711355caeac8); Time taken: 0.305 seconds
INFO : Executing command(queryId=hive_20191201155959_adcb1ad4-624e-4d85-a15c-711355caeac8): select count(*) as no_of_days from weather_final_2 where dew < 1
INFO : Query ID = hive_20191201155959_adcb1ad4-624e-4d85-a15c-711355caeac8
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_1569350662793_72365
INFO : The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_72365/
INFO : Starting Job = job_1569350662793_72365, Tracking URL = http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_72365
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_72365
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2019-12-01 15:59:14,574 Stage-1 map = 0%, reduce = 0%
INFO : 2019-12-01 15:59:20,724 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.21 sec
INFO : 2019-12-01 15:59:27,885 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6.48 sec
INFO : MapReduce Total cumulative CPU time: 6 seconds 480 msec
INFO : Ended Job = job_1569350662793_72365
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.48 sec HDFS Read: 266286 HDFS Write: 3 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 480 msec
INFO : Completed executing command(queryId=hive_20191201155959_adcb1ad4-624e-4d85-a15c-711355caeac8); Time taken: 27.467 seconds
INFO : OK
+-----+---+
| no_of_days |
+-----+---+
| 99      |
+-----+---+
1 row selected (27.935 seconds)

0: jdbc:hive2://babar.es.its.nyu.edu:10000> select year,avg(avg_temp) from weather_final_2 group by year;
INFO : Compiling command(queryId=hive_20191201160202_5174dbf0-d152-47b6-87c4-99c2baada488): select year,avg(avg_temp) from weather_final_2 group by year
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:year, type:int, comment:null), FieldSchema(name:_cl, type:double, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20191201160202_5174dbf0-d152-47b6-87c4-99c2baada488); Time taken: 0.314 seconds
INFO : Executing command(queryId=hive_20191201160202_5174dbf0-d152-47b6-87c4-99c2baada488): select year,avg(avg_temp) from weather_final_2 group by year
INFO : Query ID = hive_20191201160202_5174dbf0-d152-47b6-87c4-99c2baada488
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 not specified. Estimated from input data size: 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_1569350662793_72370
INFO : The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_72370/
INFO : Starting Job = job_1569350662793_72370, Tracking URL = http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_72370
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_72370
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2019-12-01 16:02:28,538 Stage-1 map = 0%, reduce = 0%
INFO : 2019-12-01 16:02:34,691 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.0 sec
INFO : 2019-12-01 16:02:41,861 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6.37 sec
INFO : MapReduce Total cumulative CPU time: 6 seconds 370 msec
INFO : Ended Job = job_1569350662793_72370
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.37 sec HDFS Read: 266038 HDFS Write: 94 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 370 msec
INFO : Completed executing command(queryId=hive_20191201160202_5174dbf0-d152-47b6-87c4-99c2baada488); Time taken: 35.354 seconds
INFO : OK
+-----+---+
| year |      _cl      |
+-----+---+
| 2016 | 54.18973214285711 |
| 2017 | 53.12807606263984 |
| 2018 | 53.194828957836066 |
| 2019 | 55.286327204149025 |
+-----+---+
4 rows selected (35.796 seconds)
```

```

hbr244@login-2-1:~ 
0: jdbc:hive2://babar.es.its.nyu.edu:10000> select year,avg(dew) as avg_Dew from weather_final_2 group by year;
INFO : Compiling command(queryId=hive_20191201163737_024976f3-99c4-4232-ab30-0ecddb0585db): select year,avg(dew) as avg_Dew from weather_final_2 group by year
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:year, type:int, comment:null), FieldSchema(name:avg_dew, type:double, comment:null)], properties:nu
11)
INFO : Completed compiling command(queryId=hive_20191201163737_024976f3-99c4-4232-ab30-0ecddb0585db); Time taken: 1.285 seconds
INFO : Executing command(queryId=hive_20191201163737_024976f3-99c4-4232-ab30-0ecddb0585db): select year,avg(dew) as avg_Dew from weather_final_2 group by year
INFO : Query ID = hive_20191201163737_024976f3-99c4-4232-ab30-0ecddb0585db
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 not specified. Estimated from input data size: 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_1569350662793_72439
INFO : The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_72439/
INFO : Starting Job = job_1569350662793_72439, Tracking URL = http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_72439/
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_72439
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2019-12-01 16:37:53,636 Stage-1 map = 0%, reduce = 0%
INFO : 2019-12-01 16:37:59,791 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.77 sec
INFO : 2019-12-01 16:38:05,965 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6.08 sec
INFO : MapReduce Total cumulative CPU time: 6 seconds 80 msec
INFO : Ended Job = job_1569350662793_72439
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.08 sec HDFS Read: 266111 HDFS Write: 93 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 80 msec
INFO : Completed executing command(queryId=hive_20191201163737_024976f3-99c4-4232-ab30-0ecddb0585db); Time taken: 34.764 seconds
INFO : OK
+-----+
| year | avg_dew |
+-----+
| 2016 | 42.63351934523814 |
| 2017 | 46.154399701715214 |
| 2018 | 43.44232299124904 |
| 2019 | 44.53663366336636 |
+-----+
4 rows selected (36.202 seconds)
0: jdbc:hive2://babar.es.its.nyu.edu:10000>

```

```

hbr244@login-2-1:~ 
0: jdbc:hive2://babar.es.its.nyu.edu:10000> select year,month,avg(avg_temp) as avg_temp, avg(dew) as avg_dew from weather_final_2 where month between 1 and 3 group by year,month;
INFO : Compiling command(queryId=hive_20191201163838_da132f8f-a2f3-4ca9-ad30-1a72c3afad89): select year,month,avg(avg_temp) as avg_temp, avg(dew) as avg_dew from weather_final_2 where month between 1 and 3 group by year,month
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:year, type:int, comment:null), FieldSchema(name:month, type:int, comment:null), FieldSchema(name:av
g_temp, type:double, comment:null), FieldSchema(name:avg_dew, type:double, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20191201163838_da132f8f-a2f3-4ca9-ad30-1a72c3afad89); Time taken: 1.589 seconds
INFO : Executing command(queryId=hive_20191201163838_da132f8f-a2f3-4ca9-ad30-1a72c3afad89): select year,month,avg(avg_temp) as avg_temp, avg(dew) as avg_dew from weather_final_2 where month between 1 and 3 group by year,month
INFO : Query ID = hive_20191201163838_da132f8f-a2f3-4ca9-ad30-1a72c3afad89
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 not specified. Estimated from input data size: 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_1569350662793_72440
INFO : The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_72440/
INFO : Starting Job = job_1569350662793_72440, Tracking URL = http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_72440/
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_72440
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2019-12-01 16:38:55,016 Stage-1 map = 0%, reduce = 0%
INFO : 2019-12-01 16:39:01,170 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.45 sec
INFO : 2019-12-01 16:39:08,365 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6.96 sec
INFO : MapReduce Total cumulative CPU time: 6 seconds 960 msec
INFO : Ended Job = job_1569350662793_72440
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.96 sec HDFS Read: 267536 HDFS Write: 527 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 960 msec
INFO : Completed executing command(queryId=hive_20191201163838_da132f8f-a2f3-4ca9-ad30-1a72c3afad89); Time taken: 31.69 seconds
INFO : OK
+-----+
| year | month | avg_temp | avg_dew |
+-----+
| 2016 | 1 | 32.549765258215956 | 19.922535211267604 |
| 2016 | 2 | 34.6817733990148 | 23.25418719211822 |
| 2016 | 3 | 44.43824884792627 | 30.7672811059908 |
| 2017 | 1 | 35.701612903225815 | 26.696370967741935 |
+-----+

```

```

hbr244@login-2-1:~ 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 not specified. Estimated from input data size: 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_1569350662793_72440
INFO : The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_72440/
INFO : Starting Job = job_1569350662793_72440, Tracking URL = http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_72440/
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_72440
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2019-12-01 16:38:55,016 Stage-1 map = 0%, reduce = 0%
INFO : 2019-12-01 16:39:01,170 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.45 sec
INFO : 2019-12-01 16:39:08,365 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6.96 sec
INFO : MapReduce Total cumulative CPU time: 6 seconds 960 msec
INFO : Ended Job = job_1569350662793_72440
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.96 sec HDFS Read: 267536 HDFS Write: 527 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 960 msec
INFO : Completed executing command(queryId=hive_20191201163838_da132f8f-a2f3-4ca9-ad30-1a72c3afad89); Time taken: 31.69 seconds
INFO : OK
+-----+
| year | month | avg_temp | avg_dew |
+-----+
| 2016 | 1     | 32.549765258215956 | 19.922535211267604 |
| 2016 | 2     | 34.6817733990148 | 23.25418719211822 |
| 2016 | 3     | 44.4382488479267 | 30.7672811059908 |
| 2017 | 1     | 35.701612903225815 | 26.69637096771935 |
| 2017 | 2     | 38.2545045045045 | 26.52432432432432 |
| 2017 | 3     | 36.29392712550607 | 21.577327935222673 |
| 2018 | 1     | 29.143961352657005 | 18.801932367149746 |
| 2018 | 2     | 36.435714285714305 | 29.837755102040816 |
| 2018 | 3     | 37.94838709677418 | 25.211981566820267 |
| 2019 | 1     | 30.879262672811087 | 19.48064516129031 |
| 2019 | 2     | 32.868877551020425 | 21.266836734693882 |
| 2019 | 3     | 37.91474654377881 | 24.204147465437792 |
+-----+
12 rows selected (33.484 seconds)
0: jdbc:hive2://babar.es.its.nyu.edu:10000/> 
4:41 PM 12/1/2019

0: jdbc:hive2://babar.es.its.nyu.edu:10000/> select year,avg(windspeed) from weather_final_2 group by year;
INFO : Compiling command(queryId=hive_20191201164343_2e97ee97-7017-4ef1-afa0-3ffffb9f7e1a): select year,avg(windspeed) from weather_final_2 group by year
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:year, type:int, comment:null), FieldSchema(name:c1, type:double, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20191201164343_2e97ee97-7017-4ef1-afa0-3ffffb9f7e1a); Time taken: 0.543 seconds
INFO : Executing command(queryId=hive_20191201164343_2e97ee97-7017-4ef1-afa0-3ffffb9f7e1a): select year,avg(windspeed) from weather_final_2 group by year
INFO : Query ID = hive_20191201164343_2e97ee97-7017-4ef1-afa0-3ffffb9f7e1a
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 not specified. Estimated from input data size: 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_1569350662793_72451
INFO : The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_72451/
INFO : Starting Job = job_1569350662793_72451, Tracking URL = http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_72451/
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_72451
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2019-12-01 16:43:19,564 Stage-1 map = 0%, reduce = 0%
INFO : 2019-12-01 16:43:25,710 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.83 sec
INFO : 2019-12-01 16:43:35,985 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 7.25 sec
INFO : MapReduce Total cumulative CPU time: 7 seconds 250 msec
INFO : Ended Job = job_1569350662793_72451
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 7.25 sec HDFS Read: 266040 HDFS Write: 91 SUCCESS
INFO : Total MapReduce CPU Time Spent: 7 seconds 250 msec
INFO : Completed executing command(queryId=hive_20191201164343_2e97ee97-7017-4ef1-afa0-3ffffb9f7e1a); Time taken: 32.629 seconds
INFO : OK
+-----+
| year | c1 |
+-----+
| 2016 | 6.238430059523802 |
| 2017 | 6.44780014914243 |
| 2018 | 6.323428796727136 |
| 2019 | 6.345591702027347 |
+-----+
4 rows selected (33.398 seconds)
0: jdbc:hive2://babar.es.its.nyu.edu:10000/> 
4:43 PM 12/1/2019

```

```

hbr244@login-2-1:~ 
0: jdbc:hive2://babar.es.its.nyu.edu:10000> select year,avg(visibility) as avg_visibility from weather_final_2 group by year;
INFO : Compiling command(queryId=hive_20191201164747_2c69d3df-1f54-43b8-b6f3-2aea9ae0ed67): select year,avg(visibility) as avg_visibility from weather_final_2 group by year
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:year, type:int, comment:null), FieldSchema(name:avg_visibility, type:double, comment:null)]), properties:{"year":null}
INFO : Completed compiling command(queryId=hive_20191201164747_2c69d3df-1f54-43b8-b6f3-2aea9ae0ed67); Time taken: 0.44 seconds
INFO : Executing command(queryId=hive_20191201164747_2c69d3df-1f54-43b8-b6f3-2aea9ae0ed67): select year,avg(visibility) as avg_visibility from weather_final_2 group by year
INFO : Query ID = hive_20191201164747_2c69d3df-1f54-43b8-b6f3-2aea9ae0ed67
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 not specified. Estimated from input data size: 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_1569350662793_72459
INFO : The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_72459/
INFO : Starting Job = job_1569350662793_72459, Tracking URL = http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_72459/
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_72459
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2019-12-01 16:47:31,277 Stage-1 map = 0%, reduce = 0%
INFO : 2019-12-01 16:47:36,488 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.77 sec
INFO : 2019-12-01 16:47:44,680 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 5.94 sec
INFO : MapReduce Total cumulative CPU time: 5 seconds 940 msec
INFO : Ended Job = job_1569350662793_72459
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 5.94 sec HDFS Read: 266125 HDFS Write: 93 SUCCESS
INFO : Total MapReduce CPU Time Spent: 5 seconds 940 msec
INFO : Completed executing command(queryId=hive_20191201164747_2c69d3df-1f54-43b8-b6f3-2aea9ae0ed67); Time taken: 31.521 seconds
INFO : OK
+-----+
| year | avg_visibility |
+-----+
| 2016 | 7.9807663690476165 |
| 2017 | 7.879082774049214 |
| 2018 | 7.792720763723143 |
| 2019 | 7.875766148043362 |
+-----+
4:48 PM
12/1/2019

INFO : Executing command(queryId=hive_20191201172323_2e702348-3c1f-4756-b458-a357c55d5330): select year,month,avg(avg_temp) as avg_temp, avg(dew) as avg_dew from weather_final_2 where month between 4 and 6 group by year,month
INFO : Query ID = hive_20191201172323_2e702348-3c1f-4756-b458-a357c55d5330
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 not specified. Estimated from input data size: 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_1569350662793_72527
INFO : The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_72527/
INFO : Starting Job = job_1569350662793_72527, Tracking URL = http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_72527/
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_72527
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2019-12-01 17:23:49,447 Stage-1 map = 0%, reduce = 0%
INFO : 2019-12-01 17:23:55,591 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.35 sec
INFO : 2019-12-01 17:24:03,775 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 5.6 sec
INFO : MapReduce Total cumulative CPU time: 6 seconds 560 msec
INFO : Ended Job = job_1569350662793_72527
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 5.6 sec HDFS Read: 267529 HDFS Write: 522 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 560 msec
INFO : Completed executing command(queryId=hive_20191201172323_2e702348-3c1f-4756-b458-a357c55d5330); Time taken: 34.769 seconds
INFO : OK
+-----+
| year | month | avg_temp | avg_dew |
+-----+
| 2016 | 4 | 48.035238095238135 | 33.11619047619048 |
| 2016 | 5 | 58.87880184331799 | 47.544239631336424 |
| 2016 | 6 | 68.39619047619047 | 55.347142857142885 |
| 2017 | 4 | 53.17142857142858 | 41.58781512605044 |
| 2017 | 5 | 58.34406779661011 | 47.4330508474576 |
| 2017 | 6 | 68.60000000000005 | 57.403921568627446 |
| 2018 | 4 | 45.28333333333337 | 32.85285714285713 |
| 2018 | 5 | 62.35865384615383 | 51.919230769230744 |
| 2018 | 6 | 68.10952380952384 | 57.52952380952383 |
| 2019 | 4 | 50.490430622009576 | 38.848325358851696 |
| 2019 | 5 | 58.49860465116281 | 49.27488372093024 |
+-----+
5:24 PM
12/1/2019

```

```

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 not specified. Estimated from input data size: 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_1569350662793_72527
INFO : The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_72527/
INFO : Starting Job = job_1569350662793_72527, Tracking URL = http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_72527
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p.0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_72527
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2019-12-01 17:23:49,447 Stage-1 map = 0%, reduce = 0%
INFO : 2019-12-01 17:23:58,591 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.35 sec
INFO : 2019-12-01 17:24:03,775 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6.56 sec
INFO : MapReduce Total cumulative CPU time: 6 seconds 560 msec
INFO : Ended Job = job_1569350662793_72527
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 6.56 sec   HDFS Read: 267529 HDFS Write: 522 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 560 msec
INFO : Completed executing command(queryId:hive_20191201172323_2e702348-3c1f-4756-b458-a357c55d5330); Time taken: 34.769 seconds
INFO : OK
+-----+-----+-----+-----+
| year | month | avg_temp | avg_dew |
+-----+-----+-----+-----+
| 2016 | 4     | 48.035238095238135 | 33.11619047619048 |
| 2016 | 5     | 58.87880184331799  | 47.544239631336424 |
| 2016 | 6     | 68.39619047619047  | 55.347142857142885 |
| 2017 | 4     | 53.17142857142858 | 41.58781512605044 |
| 2017 | 5     | 58.34406779661011 | 47.4330508474576  |
| 2017 | 6     | 68.60000000000005  | 57.403921568627446 |
| 2018 | 4     | 45.28333333333337 | 32.85285714285713 |
| 2018 | 5     | 62.35865384615383 | 51.919230769230744 |
| 2018 | 6     | 68.10952380952384 | 57.52952380952383 |
| 2019 | 4     | 50.490430622009576 | 38.848325358851696 |
| 2019 | 5     | 58.49860465116281 | 49.27488372093024 |
| 2019 | 6     | 68.14928229665071 | 57.15406698564592 |
+-----+-----+-----+-----+
12 rows selected (36.8 seconds)

```

```
INFO : Compiling command(queryId=hive_20191201172525_e3ee4b05-f418-452c-9de9-c736714de8c1): select year,month,avg(avg_temp) as avg_temp, avg(dew) as avg_dew from weather_final where month between 7 and 9 group by year,month
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema:[FieldSchemas:[FieldSchema(name:year, type:int, comment:null), FieldSchema(name:month, type:int, comment:null), FieldSchema(name:avg_temp, type:double, comment:null), FieldSchema(name:avg_dew, type:double, comment:null)]], properties:null
INFO : Completed compiling command(queryId=hive_20191201172525_e3ee4b05-f418-452c-9de9-c736714de8c1); Time taken: 0.44 seconds
INFO : Executing command(queryId=hive_20191201172525_e3ee4b05-f418-452c-9de9-c736714de8c1): select year,month,avg(avg_temp) as avg_temp, avg(dew) as avg_dew from weather_final where month between 7 and 9 group by year,month
INFO : Query ID = hive_20191201172525_e3ee4b05-f418-452c-9de9-c736714de8c1
INFO : Total jobs = 1
INFO : Launching Job 1 out of 1
INFO : Starting task [Stage-1:MAPPED] in serial mode
INFO : Number of reduce tasks not specified. Estimated from input data size: 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_1569350662793_72532
INFO : The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_72532/
INFO : Starting Job = job_1569350662793_72532, Tracking URL = http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_72532/
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p-0.3.0/lib/hadoop/bin/hadoop job -kill job_1569350662793_72532
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2019-12-01 17:25:32,421 Stage-1 map = 0%, reduce = 0%
INFO : 2019-12-01 17:25:39,593 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.5 sec
INFO : 2019-12-01 17:25:46,783 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6.88 sec
INFO : MapReduce Total cumulative CPU time: 6 seconds 880 msec
INFO : Ended Job = job_1569350662793_72532
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.88 sec HDFS Read: 270755 HDFS Write: 508 SUCCESS
INFO : Total Mapreduce CPU Time Spent: 6 seconds 880 msec
INFO : Completed executing command(queryId=hive_20191201172525_e3ee4b05-f418-452c-9de9-c736714de8c1); Time taken: 31.425 seconds
INFO : OK
+-----+
| year | month | avg_temp | avg_dew |
+-----+
| 2016 | 7 | 75.57205240174675 | 64.30349344978167 |
| 2016 | 8 | 76.18861788617883 | 65.67195121951224 |
| 2016 | 9 | 68.40541666666667 | 58.434583333333336 |
| 2017 | 7 | 73.68986175115208 | 64.41059907834104 |
| 2017 | 8 | 71.47142857142858 | 62.141935483870995 |
| 2017 | 9 | 67.72248803827752 | 58.831100478468905 |
+-----+
```

```

hbr244@login-2-1:~ 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 not specified. Estimated from input data size: 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_1569350662793_72532
INFO : The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_72532/
INFO : Starting Job = job_1569350662793_72532, Tracking URL = http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_72532/
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_72532
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2019-12-01 17:25:32,421 Stage-1 map = 0%, reduce = 0%
INFO : 2019-12-01 17:25:39,593 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.6 sec
INFO : 2019-12-01 17:25:46,783 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6.88 sec
INFO : MapReduce Total cumulative CPU time: 6 seconds 880 msec
INFO : Ended Job = job_1569350662793_72532
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.88 sec HDFS Read: 270755 HDFS Write: 508 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 seconds 880 msec
INFO : Completed executing command(queryId=hive_20191201172525_e3ee4b05-f418-452c-9de9-c736714de8c1); Time taken: 31.425 seconds
INFO : OK
+-----+-----+-----+
| year | month | avg_temp | avg_dew |
+-----+-----+-----+
| 2016 | 7     | 75.57205240174675 | 64.30349344978167 |
| 2016 | 8     | 76.18861788617883 | 65.67195121951224 |
| 2016 | 9     | 68.40541666666667 | 58.43458333333336 |
| 2017 | 7     | 73.68986175115208 | 64.41059907834104 |
| 2017 | 8     | 71.47142857142858 | 62.141935483870995 |
| 2017 | 9     | 67.72248803827752 | 58.831100478468905 |
| 2018 | 7     | 74.85049019607844 | 64.98431372549021 |
| 2018 | 8     | 76.07129629629627 | 68.1375 |
| 2018 | 9     | 69.05384615384614 | 62.55961538461538 |
| 2019 | 7     | 76.21013824884795 | 65.75115207373271 |
| 2019 | 8     | 72.5654377880184 | 62.94331797235024 |
| 2019 | 9     | 66.91304347826085 | 57.00676328502418 |
+-----+-----+-----+
12 rows selected (32.045 seconds)
0: jdbc:hive2://babar.es.its.nyu.edu:10000/>

```

```

hbr244@login-2-1:~ INFO : Compiling command(queryId=hive_20191201172929_db37cad9-21b9-4331-b2a5-5d956e0b6005): select year,month,avg(avg_temp) as avg_temp, avg(dew) as avg_dew from weather_final where month between 10 and 12 group by year,month
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:year, type:int, comment:null), FieldSchema(name:month, type:int, comment:null), FieldSchema(name:avg_temp, type:double, comment:null), FieldSchema(name:avg_dew, type:double, comment:null)]), properties:null
INFO : Completed compiling command(queryId=hive_20191201172929_db37cad9-21b9-4331-b2a5-5d956e0b6005); Time taken: 0.416 seconds
INFO : Executing command(queryId=hive_20191201172929_db37cad9-21b9-4331-b2a5-5d956e0b6005): select year,month,avg(avg_temp) as avg_temp, avg(dew) as avg_dew from weather_final where month between 10 and 12 group by year,month
INFO : Query ID = hive_20191201172929_db37cad9-21b9-4331-b2a5-5d956e0b6005
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 not specified. Estimated from input data size: 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_1569350662793_72550
INFO : The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_72550/
INFO : Starting Job = job_1569350662793_72550, Tracking URL = http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_72550/
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_72550
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2019-12-01 17:29:47,691 Stage-1 map = 0%, reduce = 0%
INFO : 2019-12-01 17:30:00,071 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 5.42 sec
INFO : 2019-12-01 17:30:11,352 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 10.68 sec
INFO : MapReduce Total cumulative CPU time: 10 seconds 680 msec
INFO : Ended Job = job_1569350662793_72550
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 10.68 sec HDFS Read: 270755 HDFS Write: 488 SUCCESS
INFO : Total MapReduce CPU Time Spent: 10 seconds 680 msec
INFO : Completed executing command(queryId=hive_20191201172929_db37cad9-21b9-4331-b2a5-5d956e0b6005); Time taken: 53.035 seconds
INFO : OK
+-----+-----+-----+
| year | month | avg_temp | avg_dew |
+-----+-----+-----+
| 2016 | 10    | 56.16913043478261 | 47.33217391304347 |
| 2016 | 11    | 46.484444444444435 | 35.50622222222223 |
| 2016 | 12    | 36.18749999999999 | 25.33669354838708 |
| 2017 | 10    | 61.43824884792625 | 52.13502304147466 |
| 2017 | 11    | 45.377142857142886 | 81.88809523809523 |
| 2017 | 12    | 33.388571428571428 | 22.47695852534563 |
+-----+-----+-----+

```

```

hbr244@login-2-1:~
```

INFO : Query ID = hive_20191201172929_db37cad9-21b9-4331-b2a5-5d956e0b6005
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 not specified. Estimated from input data size: 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_1569350662793_72550
INFO : The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_72550/
INFO : Starting Job = job_1569350662793_72550, Tracking URL = http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_72550
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_72550
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2019-12-01 17:29:47,691 Stage-1 map = 0%, reduce = 0%
INFO : 2019-12-01 17:30:00,071 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 5.42 sec
INFO : 2019-12-01 17:30:11,352 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 10.68 sec
INFO : MapReduce Total cumulative CPU time: 10 seconds 680 msec
INFO : Ended Job = job_1569350662793_72550
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 10.68 sec HDFS Read: 270755 HDFS Write: 488 SUCCESS
INFO : Total MapReduce CPU Time Spent: 10 seconds 680 msec
INFO : Completed executing command(queryId=hive_20191201172929_db37cad9-21b9-4331-b2a5-5d956e0b6005); Time taken: 53.035 seconds
INFO : OK
+-----+
| year | month | avg_temp | avg_dew |
+-----+
2016	10	56.16913043478261	47.33217391304347
2016	11	46.484444444444435	35.50622222222223
2016	12	36.18749999999999	25.33669354838708
2017	10	61.43824884792625	52.13502304147466
2017	11	45.377142857142886	81.88809523809523
2017	12	33.38571428571428	22.4769585234563
2018	10	55.78986175115205	47.00552995391706
2018	11	43.012254901960794	33.25931372549021
2018	12	37.80046082949312	28.807834101382483
2019	10	57.14761904761904	48.28619047619046
2019	11	54.64285714285715	41.21428571428571
+-----+
11 rows selected (53.648 seconds)
0: jdbc:hive2://babar.es.its.nyu.edu:10000/ >

```

YEAR(date1)
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name: _c0, type:int, comment:null), FieldSchema(name: _c1, type:bigint, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20191116154242_e9beac96-8f09-43b5-aebb-24664c84be99); Time taken: 0.324 seconds
INFO : Executing command(queryId=hive_20191116154242_e9beac96-8f09-43b5-aebb-24664c84be99): select YEAR(date1),count(*) from finaldata WHERE MONTH(date1) BETWEEN 4 AND 6 group by
YEAR(date1)
INFO : Query ID = hive_20191116154242_e9beac96-8f09-43b5-aebb-24664c84be99
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 not specified. Estimated from input data size: 69
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:20
INFO : Submitting tokens for job: job_1569350662793_42808
INFO : The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_42808/
INFO : Starting Job = job_1569350662793_42808, Tracking URL = http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_42808
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_42808
INFO : Hadoop job information for Stage-1: number of mappers: 20; number of reducers: 69
INFO : 2019-11-16 15:43:02.738 Stage-1 map = 0%, reduce = 0%
INFO : 2019-11-16 15:43:08.709 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 5.94 sec
INFO : 2019-11-16 15:43:09.738 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 12.26 sec
INFO : 2019-11-16 15:43:18.763 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 43.11 sec
INFO : 2019-11-16 15:43:11.787 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 175.79 sec
INFO : 2019-11-16 15:43:12.810 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 187.07 sec
INFO : 2019-11-16 15:43:16.998 Stage-1 map = 100%, reduce = 2%, Cumulative CPU 247.13 sec
INFO : 2019-11-16 15:43:18.076 Stage-1 map = 100%, reduce = 45%, Cumulative CPU 291.53 sec
INFO : 2019-11-16 15:43:19.100 Stage-1 map = 100%, reduce = 48%, Cumulative CPU 298.74 sec
INFO : 2019-11-16 15:43:20.122 Stage-1 map = 100%, reduce = 49%, Cumulative CPU 301.91 sec
INFO : 2019-11-16 15:43:21.147 Stage-1 map = 100%, reduce = 74%, Cumulative CPU 354.13 sec
INFO : 2019-11-16 15:43:22.176 Stage-1 map = 100%, reduce = 91%, Cumulative CPU 391.82 sec
INFO : 2019-11-16 15:43:23.192 Stage-1 map = 100%, reduce = 96%, Cumulative CPU 404.58 sec
INFO : 2019-11-16 15:43:24.214 Stage-1 map = 100%, reduce = 97%, Cumulative CPU 407.58 sec
INFO : 2019-11-16 15:43:25.236 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 413.28 sec
INFO : MapReduce Total cumulative CPU time: 6 minutes 53 seconds 280 msec
INFO : Ended Job = job_1569350662793_42808
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 20 Reduce: 69 Cumulative CPU: 413.28 sec HDFS Read: 4630012218 HDFS Write: 39 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 minutes 53 seconds 280 msec
INFO : Completed executing command(queryId=hive_20191116154242_e9beac96-8f09-43b5-aebb-24664c84be99); Time taken: 38.943 seconds
INFO : OK
+-----+
| _c0 | _c1 |
+-----+
| 2017 | 1978855 |
| 2018 | 2625746 |
| 2019 | 3018196 |
+-----+
3 rows selected (39.43 seconds)
```

```

[spark@babar1]
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:int, comment:null), FieldSchema(name:_c1, type:bigint, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20191116153737_ab75f133-807a-44a1-9893-5f7638117a00); Time taken: 0.45 seconds
INFO : Executing command(queryId=hive_20191116153737_ab75f133-807a-44a1-9893-5f7638117a00): select YEAR(date1),count(*) from finaldata WHERE MONTH(date1) BETWEEN 1 AND 3 group by YEAR(date1)
INFO : Query ID = hive_20191116153737_ab75f133-807a-44a1-9893-5f7638117a00
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 not specified. Estimated from input data size: 69
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:20
INFO : Submitting tokens for job: job_1569350662793_42804
INFO : The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_42804/
INFO : Starting Job = job_1569350662793_42804, Tracking URL = http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_42804/
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_42804
INFO : Hadoop job information for Stage-1: number of mappers: 20; number of reducers: 69
INFO : 2019-11-16 15:37:14,801 Stage-1 map = 0%, reduce = 0%
INFO : 2019-11-16 15:37:20,955 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 8.46 sec
INFO : 2019-11-16 15:37:21,988 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 14.79 sec
INFO : 2019-11-16 15:37:23,003 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 40.67 sec
INFO : 2019-11-16 15:37:24,033 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 157.86 sec
INFO : 2019-11-16 15:37:25,058 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 168.57 sec
INFO : 2019-11-16 15:37:30,175 Stage-1 map = 100%, reduce = 29%, Cumulative CPU 231.21 sec
INFO : 2019-11-16 15:37:31,200 Stage-1 map = 100%, reduce = 48%, Cumulative CPU 272.36 sec
INFO : 2019-11-16 15:37:33,247 Stage-1 map = 100%, reduce = 77%, Cumulative CPU 336.29 sec
INFO : 2019-11-16 15:37:35,294 Stage-1 map = 100%, reduce = 94%, Cumulative CPU 376.64 sec
INFO : 2019-11-16 15:37:36,317 Stage-1 map = 100%, reduce = 97%, Cumulative CPU 387.21 sec
INFO : 2019-11-16 15:37:37,339 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 393.44 sec
INFO : MapReduce Total cumulative CPU time: 6 minutes 33 seconds 440 msec
INFO : Ended Job = job_1569350662793_42804
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 20 Reduce: 69 Cumulative CPU: 393.44 sec HDFS Read: 4630012218 HDFS Write: 38 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 minutes 33 seconds 440 msec
INFO : Completed executing command(queryId=hive_20191116153737_ab75f133-807a-44a1-9893-5f7638117a00); Time taken: 37.785 seconds
INFO : OK
+-----+-----+
| _c0 | _c1 |
+-----+-----+
| 2017 | 889024 |
| 2018 | 1053216 |
| 2019 | 1348218 |
+-----+-----+
3 rows selected (38.384 seconds)
0: jdbc:hive2://babar.es.its.nyu.edu:10000/> ■

```

```

5 rows selected (41.272 seconds)
0: jdbc:hive2://babar.es.its.nyu.edu:10000/> select YEAR(date1), COUNT(*) from finaldata group by YEAR(date1);
INFO : Compiling command(queryId=hive_20191116145555_889ba5f0-7b51-4365-9ca9-c908db9b6fb4): select YEAR(date1), COUNT(*) from finaldata group by YEAR(date1)
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:int, comment:null), FieldSchema(name:_c1, type:bigint, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20191116145555_889ba5f0-7b51-4365-9ca9-c908db9b6fb4); Time taken: 0.367 seconds
INFO : Executing command(queryId=hive_20191116145555_889ba5f0-7b51-4365-9ca9-c908db9b6fb4): select YEAR(date1), COUNT(*) from finaldata group by YEAR(date1)
INFO : Query ID = hive_20191116145555_889ba5f0-7b51-4365-9ca9-c908db9b6fb4
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 not specified. Estimated from input data size: 69
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:20
INFO : Submitting tokens for job: job_1569350662793_42764
INFO : The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_42764/
INFO : Starting Job = job_1569350662793_42764, Tracking URL = http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_42764/
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_42764
INFO : Hadoop job information for Stage-1: number of mappers: 20; number of reducers: 69
INFO : 2019-11-16 14:55:17,088 Stage-1 map = 0%, reduce = 0%
INFO : 2019-11-16 14:55:24,843 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 22.87 sec
INFO : 2019-11-16 14:55:25,864 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 93.94 sec
INFO : 2019-11-16 14:55:26,885 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 124.18 sec
INFO : 2019-11-16 14:55:27,909 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 132.98 sec
INFO : 2019-11-16 14:55:30,991 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 140.44 sec
INFO : 2019-11-16 14:55:32,013 Stage-1 map = 100%, reduce = 29%, Cumulative CPU 200.37 sec
INFO : 2019-11-16 14:55:34,056 Stage-1 map = 100%, reduce = 45%, Cumulative CPU 237.35 sec
INFO : 2019-11-16 14:55:35,078 Stage-1 map = 100%, reduce = 48%, Cumulative CPU 244.27 sec
INFO : 2019-11-16 14:55:36,099 Stage-1 map = 100%, reduce = 77%, Cumulative CPU 309.11 sec
INFO : 2019-11-16 14:55:38,141 Stage-1 map = 100%, reduce = 93%, Cumulative CPU 345.31 sec
INFO : 2019-11-16 14:55:39,162 Stage-1 map = 100%, reduce = 97%, Cumulative CPU 355.73 sec
INFO : 2019-11-16 14:55:48,182 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 361.39 sec
INFO : MapReduce Total cumulative CPU time: 6 minutes 1 seconds 390 msec
INFO : Ended Job = job_1569350662793_42764
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 20 Reduce: 69 Cumulative CPU: 361.39 sec HDFS Read: 4629997751 HDFS Write: 62 SUCCESS
INFO : Total MapReduce CPU Time Spent: 6 minutes 1 seconds 390 msec
INFO : Completed executing command(queryId=hive_20191116145555_889ba5f0-7b51-4365-9ca9-c908db9b6fb4); Time taken: 39.802 seconds
INFO : OK
+-----+-----+
| _c0 | _c1 |
+-----+-----+
| NULL | 704355 |
| 2016 | 1512401 |
| 2017 | 6954513 |
| 2018 | 8573274 |
| 2019 | 8077645 |
+-----+-----+
5 rows selected (41.272 seconds)
0: jdbc:hive2://babar.es.its.nyu.edu:10000/> ■

```

```
error. Error while compiling statement: FAILED: ParseException line 2:7 missing EOF at 'from' near '*' (state=42000,code=40000)
0: jdbc:hive2://babar.es.its.nyu.edu:10000> insert overwrite directory '/user/mn2643/FinalProject/hiveoutput/'
[...]
Error: Error while compiling statement: FAILED: SemanticException Error creating temporary folder on: /user/mn2643/FinalProject/hiveoutput (state=42000,code=40000)
0: jdbc:hive2://babar.es.its.nyu.edu:10000>
```

```
[mn2643@login-2-1 HiveOutput]$ hdfs dfs -ls /user/mn2643/hivefinal/
Found 20 items
-rw-r----+ 3 hive users 411489551 2019-11-16 13:34 /user/mn2643/hivefinal/000000_0
-rw-r----+ 3 hive users 279925104 2019-11-16 13:34 /user/mn2643/hivefinal/000001_0
-rw-r----+ 3 hive users 274035484 2019-11-16 13:34 /user/mn2643/hivefinal/000002_0
-rw-r----+ 3 hive users 274388200 2019-11-16 13:34 /user/mn2643/hivefinal/000003_0
-rw-r----+ 3 hive users 274220045 2019-11-16 13:34 /user/mn2643/hivefinal/000004_0
-rw-r----+ 3 hive users 273122387 2019-11-16 13:34 /user/mn2643/hivefinal/000005_0
-rw-r----+ 3 hive users 272868060 2019-11-16 13:34 /user/mn2643/hivefinal/000006_0
-rw-r----+ 3 hive users 272424041 2019-11-16 13:34 /user/mn2643/hivefinal/000007_0
-rw-r----+ 3 hive users 272734869 2019-11-16 13:34 /user/mn2643/hivefinal/000008_0
-rw-r----+ 3 hive users 272281524 2019-11-16 13:34 /user/mn2643/hivefinal/000009_0
-rw-r----+ 3 hive users 271711677 2019-11-16 13:34 /user/mn2643/hivefinal/000010_0
-rw-r----+ 3 hive users 271402259 2019-11-16 13:34 /user/mn2643/hivefinal/000011_0
-rw-r----+ 3 hive users 271474858 2019-11-16 13:34 /user/mn2643/hivefinal/000012_0
-rw-r----+ 3 hive users 271065040 2019-11-16 13:34 /user/mn2643/hivefinal/000013_0
-rw-r----+ 3 hive users 268067557 2019-11-16 13:34 /user/mn2643/hivefinal/000014_0
-rw-r----+ 3 hive users 136335306 2019-11-16 13:34 /user/mn2643/hivefinal/000015_0
-rw-r----+ 3 hive users 135530147 2019-11-16 13:34 /user/mn2643/hivefinal/000016_0
-rw-r----+ 3 hive users 134993632 2019-11-16 13:34 /user/mn2643/hivefinal/000017_0
-rw-r----+ 3 hive users 130774779 2019-11-16 13:34 /user/mn2643/hivefinal/000018_0
-rw-r----+ 3 hive users 125677399 2019-11-16 13:34 /user/mn2643/hivefinal/000019_0
```

areazipcode	totalaccidents	totalstations	totaltrips
11211	42120	30	386485
10002	26280	24	1046098
10009	21408	32	1298258
11206	20988	22	113445
11217	19964	23	235992
11238	18600	25	289167
10011	17380	20	1158090
11215	14732	29	309464
11201	14252	28	485190
10003	14098	19	988688
11222	13832	28	240994
10001	11340	18	858031
11237	10738	13	56496
11231	10280	40	398846
10019	9960	30	1461734
11205	8976	12	165131
11221	8400	21	224021
11216	8359	13	106297
10029	7144	19	191730
10017	6692	14	747498
11101	6272	32	226525
10022	5742	18	597315
11207	5742	9	113665
11225	5512	13	54757
10023	5421	13	687025
10012	5310	9	528001
11106	4959	19	59676
10028	4953	13	242603
10018	4698	9	413472
10021	4511	13	362006
11102	4464	16	49090
10035	4386	17	213128
10026	3360	7	130536
10027	3185	13	98435
10025	3059	19	492391
10014	2754	18	963819
10016	2665	13	640069
11213	2515	5	101805
10013	2119	13	544069
10037	2030	10	533589
10065	1980	6	185239
10038	1568	14	419346
11103	1536	3	10589
11233	1359	3	36520
11220	1270	2	632
11105	987	7	31011
10007	924	7	335448
10199	780	3	118532
11369	664	2	10346
11208	654	2	41574
10005	584	8	407613
11109	395	5	58693

```

INFO : number of splits:20
INFO : Submitting tokens for job: job_1569350662793_43033
INFO : The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_43033/
INFO : Starting Job = job_1569350662793_43033, Tracking URL = http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_43033/
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_43033
INFO : Hadoop job information for Stage-2: number of mappers: 20; number of reducers: 69
INFO : 2019-11-16 20:10:15,863 Stage-2 map = 0%, reduce = 0%
INFO : 2019-11-16 20:10:22,099 Stage-2 map = 5%, reduce = 0%, Cumulative CPU 6.67 sec
INFO : 2019-11-16 20:10:23,123 Stage-2 map = 10%, reduce = 0%, Cumulative CPU 13.25 sec
INFO : 2019-11-16 20:10:24,148 Stage-2 map = 50%, reduce = 0%, Cumulative CPU 72.3 sec
INFO : 2019-11-16 20:10:25,178 Stage-2 map = 95%, reduce = 0%, Cumulative CPU 143.12 sec
INFO : 2019-11-16 20:10:31,322 Stage-2 map = 100%, reduce = 0%, Cumulative CPU 152.93 sec
INFO : 2019-11-16 20:10:32,345 Stage-2 map = 100%, reduce = 78%, Cumulative CPU 357.35 sec
INFO : 2019-11-16 20:10:33,368 Stage-2 map = 100%, reduce = 97%, Cumulative CPU 408.28 sec
INFO : 2019-11-16 20:10:36,434 Stage-2 map = 100%, reduce = 100%, Cumulative CPU 421.42 sec
INFO : MapReduce Total cumulative CPU time: 7 minutes 1 seconds 420 msec
INFO : Ended Job = job_1569350662793_43033
INFO : Launching Job 2 out of 2
INFO : Starting task [Stage-3: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:22
INFO : Submitting tokens for job: job_1569350662793_43035
INFO : The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_43035/
INFO : Starting Job = job_1569350662793_43035, Tracking URL = http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_43035/
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_43035
INFO : Hadoop job information for Stage-3: number of mappers: 22; number of reducers: 1
INFO : 2019-11-16 20:10:53,578 Stage-3 map = 0%, reduce = 0%
INFO : 2019-11-16 20:10:57,685 Stage-3 map = 5%, reduce = 0%, Cumulative CPU 1.56 sec
INFO : 2019-11-16 20:10:59,733 Stage-3 map = 100%, reduce = 0%, Cumulative CPU 34.17 sec
INFO : 2019-11-16 20:11:04,846 Stage-3 map = 100%, reduce = 100%, Cumulative CPU 37.48 sec
INFO : MapReduce Total cumulative CPU time: 37 seconds 488 msec
INFO : Ended Job = job_1569350662793_43035
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-2: Map: 20 Reduce: 69 Cumulative CPU: 421.42 sec HDFS Read: 4630045759 HDFS Write: 8471 SUCCESS
INFO : Stage-Stage-3: Map: 22 Reduce: 1 Cumulative CPU: 37.48 sec HDFS Read: 69736 HDFS Write: 45 SUCCESS
INFO : Total MapReduce CPU Time Spent: 7 minutes 38 seconds 900 msec
INFO : Completed executing command(queryId:hive_20191116200909_0a8ae438-ac79-4bb9-847a-a5446c779dd2); Time taken: 67.186 seconds
INFO : OK
+-----+-----+
| postcode.zipcode | total |
+-----+-----+
| 11231 | 40 |
| 11101 | 32 |
| 10009 | 32 |
| 10019 | 30 |
| 11211 | 30 |
+-----+-----+
5 rows selected (67.713 seconds)

```