

Hive Scripts and Result-Screenshots

1. Create Hive table and some basic query operations.

1. First, create the Hive table using the output txt file. (I first created a test file).

```
create external table test2 (weekName string,timePeriod int,address string,lon
double,lat double,avgVolume double) row format delimited fields terminated by
',' location '/user/xf569/CBDA/test2-output/part-r-00000';
```

2. Some trial commands

```
show tables;
```

```
0: jdbc:hive2://babar.es.its.nyu.edu:10000/> show tables;
INFO : Compiling command(queryId=hive_20191110191818_88ab19e6-5420-4c98-ad60-8e5bc0f242cb): show tables
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:tab_name, type:string, comment:from deserializer)], properties:null)
INFO : Completed compiling command(queryId=hive_20191110191818_88ab19e6-5420-4c98-ad60-8e5bc0f242cb); Time taken: 0.063 seconds
INFO : Executing command(queryId=hive_20191110191818_88ab19e6-5420-4c98-ad60-8e5bc0f242cb): show tables
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command(queryId=hive_20191110191818_88ab19e6-5420-4c98-ad60-8e5bc0f242cb); Time taken: 0.135 seconds
INFO : OK
+-----+
| tab_name |
+-----+
| test2    |
| w1       |
| w10      |
| w11      |
| w3       |
| w30      |
+-----+
6 rows selected (0.205 seconds)
```

```
describe test2;
```

```
+-----+-----+-----+
| col_name | data_type | comment |
+-----+-----+-----+
| weekname | string    |         |
| timeperiod | int       |         |
| address   | string    |         |
| lon       | double    |         |
| lat       | double    |         |
| avgvolume | double    |         |
+-----+-----+-----+
```

```
select * from test2;
```

```
+-----+-----+-----+-----+-----+-----+
| test2.weekname | test2.timeperiod | test2.address | test2.lon | test2.lat | test2.avgvolume |
+-----+-----+-----+-----+-----+-----+
| 1               | 0                 | 1 Ave & E 16 St | 40.73221853 | -73.98165557 | 5.5             |
| 1               | 3                 | 8 Ave & W 31 St | 40.7505853470215 | -73.9946848154068 | 2.0             |
| 2               | 10                | E 3 St & 1 Ave | 40.72467721 | -73.98783413 | 2.0             |
| 4               | 0                 | 1 Ave & E 16 St | 40.73221853 | -73.98165557 | 5.0             |
| 4               | 3                 | 8 Ave & W 31 St | 40.7505853470215 | -73.9946848154068 | 4.0             |
+-----+-----+-----+-----+-----+-----+
5 rows selected (0.77 seconds)
```

```
select * from test2 where weekname = "2" and timeperiod = 10;
```

```

0: jdbc:hive2://babar.es.its.nyu.edu:10000/> select * from test2 where weekname = "2" and timeperiod = 10;
Error: Error while compiling statement: FAILED: ParseException line 1:40 missing EOF at ',' near '"2"' (state=42000,code=40000)
0: jdbc:hive2://babar.es.its.nyu.edu:10000/> select * from test2 where (weekname = "2" and timeperiod = 10);
Error: Error while compiling statement: FAILED: ParseException line 1:41 mismatched input ',' expecting ')' near '"2"' in expression specification (state=42000,code=40000)
0: jdbc:hive2://babar.es.its.nyu.edu:10000/> select * from test2 where weekname = "2" and timeperiod = 10;
Error: Error while compiling statement: FAILED: ParseException line 1:42 mismatched input 'timeperiod' expecting ')' near '"2"' in expression specification (state=42000,code=40000)
0: jdbc:hive2://babar.es.its.nyu.edu:10000/> select * from test2 where weekname = "2" and timeperiod = 10;
Error: Error while compiling statement: FAILED: ParseException line 1:41 missing EOF at 'timeperiod' near '"2"' (state=42000,code=40000)
0: jdbc:hive2://babar.es.its.nyu.edu:10000/> select * from test2 where weekname = "2" and timeperiod = 10;
INFO : Compiling command(queryId=hive_20191110192626_670800bb-2371-4c72-9fb9-58ef2a2e9de8): select * from test2 where weekname = "2" and timeperiod = 10
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name=test2.weekname, type:string, comment:null), FieldSchema(name=test2.timeperiod, type:int, comment:null), FieldSchema(name=test2.address, type:string, comment:null), FieldSchema(name=test2.lon, type:double, comment:null), FieldSchema(name=test2.lat, type:double, comment:null), FieldSchema(name=test2.avgvolume, type:double, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20191110192626_670800bb-2371-4c72-9fb9-58ef2a2e9de8): Time taken: 0.488 seconds
INFO : Executing command(queryId=hive_20191110192626_670800bb-2371-4c72-9fb9-58ef2a2e9de8): select * from test2 where weekname = "2" and timeperiod = 10
INFO : Query ID = hive_20191110192626_670800bb-2371-4c72-9fb9-58ef2a2e9de8
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 is set to 0 since there's no reduce operator
INFO : number of splits:1
INFO : Submitting tokens for job: job_1569350662793_37987
INFO : The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_37987/
INFO : Starting Job = job_1569350662793_37987, Tracking URL = http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_37987/
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_37987
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 0
INFO : 2019-11-10 19:26:59.014 Stage-1 map = 0%, reduce = 0%
INFO : 2019-11-10 19:27:05.317 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.82 sec
INFO : MapReduce Total cumulative CPU time: 2 seconds 820 msec
INFO : Ended Job = job_1569350662793_37987
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Cumulative CPU: 2.82 sec HDFS Read: 5320 HDFS Write: 49 SUCCESS
INFO : Total MapReduce CPU Time Spent: 2 seconds 820 msec
INFO : Completed executing command(queryId=hive_20191110192626_670800bb-2371-4c72-9fb9-58ef2a2e9de8): Time taken: 27.839 seconds
INFO : OK
+-----+-----+-----+-----+-----+-----+
| test2.weekname | test2.timeperiod | test2.address | test2.lon | test2.lat | test2.avgvolume |
+-----+-----+-----+-----+-----+-----+
| 2 | 10 | E 3 St & 1 Ave | 40.72467721 | -73.98783413 | 2.0 |
+-----+-----+-----+-----+-----+
1 row selected (28.461 seconds)

```

3. Create the final table for all data.

In Hive command environment:

```

create external table average1 (weekName string,timePeriod int,address
string,lon double,lat double,avgVolume double) row format delimited fields
terminated by ',' location '/user/xf569/CBDA/average-output/';

```

```

0: jdbc:hive2://babar.es.its.nyu.edu:10000/> create external table average1 (weekName string,timePeriod int,address string,lon double,lat double,avgVolume double) row format
delimited fields terminated by ',' location '/user/xf569/CBDA/average-output/';
INFO : Compiling command(queryId=hive_20191110194343_bb00e188-2373-49f8-883a-ab940ed462c9): create external table average1 (weekName string,timePeriod int,address string,lon
double,lat double,avgVolume double) row format delimited fields terminated by ',' location '/user/xf569/CBDA/average-output/'
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name=average1.weekname, type:string, comment:null), FieldSchema(name=average1.timeperiod, type:int, comment:n
ull), FieldSchema(name=average1.address, type:string, comment:null), FieldSchema(name=average1.lon, type:double, comment:null), FieldSchema(name=average1.lat, type:double, c
omment:null), FieldSchema(name=average1.avgvolume, type:double, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20191110194343_bb00e188-2373-49f8-883a-ab940ed462c9): create external table average1 (weekName string,timePeriod int,address string,lon
double,lat double,avgVolume double) row format delimited fields terminated by ',' location '/user/xf569/CBDA/average-output/'
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command(queryId=hive_20191110194343_bb00e188-2373-49f8-883a-ab940ed462c9): Time taken: 0.666 seconds
INFO : OK
No rows affected (0.911 seconds)

```

```
select * from average1 limit 20;
```

```

0: jdbc:hive2://babar.es.its.nyu.edu:10000/> select * from average1 limit 20;
INFO : Compiling command(queryId=hive_20191110194444_672692b2-0d27-474c-932b-160e47f69a79): select * from average1 limit 20
INFO : Semantic Analysis Completed
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name=average1.weekname, type:string, comment:null), FieldSchema(name=average1.timeperiod, type:int, comment:n
ull), FieldSchema(name=average1.address, type:string, comment:null), FieldSchema(name=average1.lon, type:double, comment:null), FieldSchema(name=average1.lat, type:double, c
omment:null), FieldSchema(name=average1.avgvolume, type:double, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20191110194444_672692b2-0d27-474c-932b-160e47f69a79): Time taken: 0.981 seconds
INFO : Executing command(queryId=hive_20191110194444_672692b2-0d27-474c-932b-160e47f69a79): select * from average1 limit 20
INFO : Completed executing command(queryId=hive_20191110194444_672692b2-0d27-474c-932b-160e47f69a79): Time taken: 0.004 seconds
INFO : OK
+-----+-----+-----+-----+-----+-----+
| average1.weekname | average1.timeperiod | average1.address | average1.lon | average1.lat | average1.avgvolume |
+-----+-----+-----+-----+-----+-----+
| 1 | 0 | 1 Ave & E 110 St | 40.7923272 | -73.9383 | 1.0 |
| 1 | 0 | 1 Ave & E 16 St | 40.73221853 | -73.98165557 | 5.75 |
| 1 | 0 | 1 Ave & E 18 St | 40.731812191966315 | -73.98054420948029 | 3.6666666666666665 |
| 1 | 0 | 1 Ave & E 30 St | 40.74144387 | -73.97536082 | 1.3333333333333333 |
| 1 | 0 | 1 Ave & E 44 St | 40.75001986 | -73.96905301 | 2.0 |
| 1 | 0 | 1 Ave & E 62 St | 40.7612274 | -73.96094022 | 2.0 |
| 1 | 0 | 1 Ave & E 78 St | 40.77140426 | -73.9535166 | 1.3333333333333333 |
| 1 | 0 | 10 St & 5 Ave | 40.6686273 | -73.98700053 | 1.0 |
| 1 | 0 | 10 St & 7 Ave | 40.6662078 | -73.98199886 | 1.3333333333333333 |
| 1 | 0 | 11 Ave & W 27 St | 40.751306 | -74.005226 | 3.25 |
| 1 | 0 | 11 Ave & W 41 St | 40.76030096 | -73.99884222 | 2.0 |
| 1 | 0 | 11 Ave & W 59 St | 40.771522 | -73.990541 | 1.3333333333333333 |
| 1 | 0 | 12 Ave & W 40 St | 40.76087502 | -74.00277668 | 2.0 |
| 1 | 0 | 12 St & 4 Ave | 40.668603 | -73.9904394 | 1.0 |
| 1 | 0 | 14 St & 5 Ave | 40.666207 | -73.98095053 | 1.0 |
| 1 | 0 | 14 St & 7 Ave | 40.663779 | -73.98396846 | 1.0 |
| 1 | 0 | 19 St & 24 Ave | 40.7774552 | -73.9225403 | 1.0 |
| 1 | 0 | 2 Ave & E 104 St | 40.7892105 | -73.94370784 | 1.0 |
| 1 | 0 | 2 Ave & E 122 St | 40.8006721 | -73.9349 | 1.0 |
| 1 | 0 | 2 Ave & E 31 St | 40.74290902 | -73.97706058 | 1.75 |
+-----+-----+-----+-----+-----+-----+
10 rows selected (1.069 seconds)

```

2. Hive analysis for our project

1. Track the Commuting Time: get insights of NYC office area.

Sample: Mean Value of Weekdays (Mon-Fri) From 7 AM-9 AM:

HiveQL:

```
FROM average1 SELECT address,lon,lat,(sum(avgvolume)/5.0) AS meanAvg WHERE
((weekname != "6" and weekname != "7") and (timeperiod = 7 or timeperiod = 8))
GROUP BY address,lon,lat SORT by meanAvg desc limit 10;
```

It might take a long time to run.

Screenshot of the result:

```
INFO : 2019-11-30 10:40:26.515 Stage-2 map = 100%, reduce = 100%, Cumulative CPU 5.19 sec
INFO : MapReduce Total cumulative CPU time: 5 seconds 190 msec
INFO : Ended Job = job_1569350662793_69680
INFO : Launching Job 3 out of 3
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:1
INFO : Submitting tokens for job: job_1569350662793_69682
INFO : The url to track the job: http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_69682/
INFO : Starting Job = job_1569350662793_69682, Tracking URL = http://babar.es.its.nyu.edu:8088/proxy/application_1569350662793_69682/
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_69682
INFO : Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 1
INFO : 2019-11-30 10:40:47.111 Stage-3 map = 0%, reduce = 0%
INFO : 2019-11-30 10:40:52.259 Stage-3 map = 100%, reduce = 0%, Cumulative CPU 1.76 sec
INFO : 2019-11-30 10:40:59.428 Stage-3 map = 100%, reduce = 100%, Cumulative CPU 5.15 sec
INFO : MapReduce Total cumulative CPU time: 5 seconds 150 msec
INFO : Ended Job = job_1569350662793_69682
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 11.98 sec HDFS Read: 5493930 HDFS Write: 46736 SUCCESS
INFO : Stage-Stage-2: Map: 1 Reduce: 1 Cumulative CPU: 5.19 sec HDFS Read: 51234 HDFS Write: 699 SUCCESS
INFO : Stage-Stage-3: Map: 1 Reduce: 1 Cumulative CPU: 5.15 sec HDFS Read: 6061 HDFS Write: 608 SUCCESS
INFO : Total MapReduce CPU Time Spent: 22 seconds 320 msec
INFO : Completed executing command(queryId=hive_20191130103838_ad6081f8-7060-444c-bda7-3efd085b7996): Time taken: 142.663 seconds
INFO : OK
+-----+-----+-----+-----+
| address | lon | lat | meanavg |
+-----+-----+-----+-----+
| Pershing Square North | 40.751873 | -73.977706 | 121.33000000000001 |
| W 52 St & 6 Ave | 40.76132983124814 | -73.97982001304626 | 108.73333333333332 |
| Grand Army Plaza & Central Park S | 40.7643971 | -73.97371465 | 84.59 |
| W 44 St & 5 Ave | 40.75508254 | -73.98014437 | 74.38 |
| W 52 St & 5 Ave | 40.75992262 | -73.97648516 | 74.2 |
| Broadway & E 22 St | 40.7403432 | -73.98955109 | 61.84333333333333 |
| West St & Chambers St | 40.71754834 | -74.01322060 | 60.92999999999999 |
| E 55 St & 3 Ave | 40.758996559605116 | -73.96865397691727 | 50.26333333333333 |
| W 21 St & 6 Ave | 40.74173969 | -73.99415556 | 49.19333333333334 |
| E 48 St & 5 Ave | 40.75724567911726 | -73.97805914282799 | 48.47 |
+-----+-----+-----+-----+
10 rows selected (144.813 seconds)
0: jdbc:hive2://babar.es.its.nyu.edu:10000/
```

2. Track the Friday night: get insights of Nightlife/Entertainment are in NYC.

Sample: Friday Night 8PM-9PM

```
select * from average1 where (weekname = "5" and timeperiod = 20) sort by
avgvolume desc limit 10;
```

```

INFO : Starting Job = job_1569350662793_69694, Tracking URL = http://babar.es.its.nyu.edu:8080/proxy/application_1569350662793_69694/
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_69694
INFO : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO : 2019-11-30 10:58:52,589 Stage-1 map = 0%, reduce = 0%
INFO : 2019-11-30 10:58:58,722 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.73 sec
INFO : 2019-11-30 10:59:05,897 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_1569350662793_69694
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_69695
INFO : The url to track the job: http://babar.es.its.nyu.edu:8080/proxy/application_1569350662793_69695/
INFO : Starting Job = job_1569350662793_69695, Tracking URL = http://babar.es.its.nyu.edu:8080/proxy/application_1569350662793_69695/
INFO : Kill Command = /opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/lib/hadoop/bin/hadoop job -kill job_1569350662793_69695
INFO : Hadoop job information for Stage-2: number of mappers: 1; number of reducers: 1
INFO : 2019-11-30 10:59:23,684 Stage-2 map = 0%, reduce = 0%
INFO : 2019-11-30 10:59:28,825 Stage-2 map = 100%, reduce = 0%, Cumulative CPU 1.5 sec
INFO : 2019-11-30 10:59:36,016 Stage-2 map = 100%, reduce = 100%, Cumulative CPU 5.31 sec
INFO : MapReduce Total cumulative CPU time: 5 seconds 310 msec
INFO : Ended Job = job_1569350662793_69695
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 7.37 sec HDFS Read: 5492198 HDFS Write: 707 SUCCESS
INFO : Stage-Stage-2: Map: 1 Reduce: 1 Cumulative CPU: 5.31 sec HDFS Read: 6634 HDFS Write: 567 SUCCESS
INFO : Total MapReduce CPU Time Spent: 12 seconds 680 msec
INFO : Completed executing command(queryId=hive_20191130105858_ac87076f-8bec-4f14-ac9a-93ec0d13f29e): Time taken: 61.528 seconds
INFO : OK
+-----+
| averagel.weekname | averagel.timeperiod | averagel.address | averagel.lon | averagel.lat | averagel.avgvolume |
+-----+
| 5 | 20 | E 7 St & Avenue A | 40.72621788 | -73.98379855 | 40.0 |
| 5 | 20 | E 39 St & 2 Ave | 40.74780373 | -73.9734419 | 18.0 |
| 5 | 20 | E 12 St & Avenue B | 40.728048571362606 | -73.97881150245667 | 17.0 |
| 5 | 20 | W 21 St & 6 Ave | 40.74173969 | -73.99415556 | 17.0 |
| 5 | 20 | Cooper Square & Astor Pl | 40.72951496224949 | -73.99075269699897 | 16.0 |
| 5 | 20 | Pershing Square North | 40.751873 | -73.977786 | 15.6 |
| 5 | 20 | LaGuardia Pl & W 3 St | 40.72917025 | -73.99810231 | 15.5 |
| 5 | 20 | Carmine St & 6 Ave | 40.73038599 | -74.00214988 | 15.0 |
| 5 | 20 | 1 Ave & E 16 St | 40.73221853 | -73.98165557 | 14.6 |
| 5 | 20 | Greenwich St & W Houston St | 40.728846 | -74.008591 | 14.4 |
+-----+
10 rows selected (62.191 seconds)
0: jdbc:hive2://babar.es.its.nyu.edu:10000/>

```

- Track the Sunday afternoon: get insights of where do New Yorkers spend the weekend afternoon.

Sample: Sunday 2PM-3PM

```

select * from averagel1 where (weekname = "7" and timeperiod = 14) sort by
avgvolume desc limit 10;

```

```

+-----+
| averagel.weekname | averagel.timeperiod | averagel.address | averagel.lon | averagel.lat | averagel.avgvolume |
+-----+
| 7 | 14 | Broadway & E 14 St | 40.73454567 | -73.99074142 | 28.5 |
| 7 | 14 | E 17 St & Broadway | 40.73704984 | -73.99089296 | 28.5 |
| 7 | 14 | Columbus Ave & W 72 St | 40.7770575 | -73.97898475 | 24.5 |
| 7 | 14 | Broadway & W 58 St | 40.76695317 | -73.98169333 | 21.75 |
| 7 | 14 | W 21 St & 6 Ave | 40.74173969 | -73.99415556 | 20.75 |
| 7 | 14 | Cleveland Pl & Spring St | 40.722103786686034 | -73.99724900722504 | 20.0 |
| 7 | 14 | University Pl & E 14 St | 40.73492695 | -73.99200509 | 19.5 |
| 7 | 14 | Grand Army Plaza & Central Park S | 40.7643971 | -73.97371465 | 18.75 |
| 7 | 14 | W 22 St & 10 Ave | 40.74691959 | -74.00451887 | 18.25 |
| 7 | 14 | Central Park S & 6 Ave | 40.76590936 | -73.97634151 | 17.666666666666668 |
+-----+
10 rows selected (57.247 seconds)
0: jdbc:hive2://babar.es.its.nyu.edu:10000/>

```

- Track the volume changing trends on a stable location on Monday (choose the place we find has most records on commuting time in step 1)

```

select * from averagel1 where (weekname = "1" and address = "Pershing Square
North") order by cast(timeperiod as int) asc;

```

```

INFO : Ended Job = job_1569350662793_69716
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 7.9 sec HDFS Read: 5492935 HDFS Write: 1213 SUCCESS
INFO : Total MapReduce CPU Time Spent: 7 seconds 900 msec
INFO : Completed executing command(queryId=hive_20191130111717_14e33fbf-efa5-4cbe-9f28-2accc1be4b58): Time taken: 32.448 seconds
INFO : OK
+-----+-----+-----+-----+-----+-----+
| average1.weekname | average1.timeperiod | average1.address | average1.lon | average1.lat | average1.avgvolume |
+-----+-----+-----+-----+-----+-----+
| 1 | 0 | Pershing Square North | 40.751873 | -73.977706 | 1.25 |
| 1 | 1 | Pershing Square North | 40.751873 | -73.977706 | 2.3333333333333335 |
| 1 | 3 | Pershing Square North | 40.751873 | -73.977706 | 1.0 |
| 1 | 4 | Pershing Square North | 40.751873 | -73.977706 | 1.0 |
| 1 | 5 | Pershing Square North | 40.751873 | -73.977706 | 3.0 |
| 1 | 6 | Pershing Square North | 40.751873 | -73.977706 | 3.5 |
| 1 | 7 | Pershing Square North | 40.751873 | -73.977706 | 2.0 |
| 1 | 8 | Pershing Square North | 40.751873 | -73.977706 | 9.25 |
| 1 | 9 | Pershing Square North | 40.751873 | -73.977706 | 9.25 |
| 1 | 10 | Pershing Square North | 40.751873 | -73.977706 | 13.25 |
| 1 | 11 | Pershing Square North | 40.751873 | -73.977706 | 14.5 |
| 1 | 12 | Pershing Square North | 40.751873 | -73.977706 | 9.5 |
| 1 | 13 | Pershing Square North | 40.751873 | -73.977706 | 9.5 |
| 1 | 14 | Pershing Square North | 40.751873 | -73.977706 | 8.5 |
| 1 | 15 | Pershing Square North | 40.751873 | -73.977706 | 12.5 |
| 1 | 16 | Pershing Square North | 40.751873 | -73.977706 | 9.5 |
| 1 | 17 | Pershing Square North | 40.751873 | -73.977706 | 9.75 |
| 1 | 18 | Pershing Square North | 40.751873 | -73.977706 | 7.75 |
| 1 | 19 | Pershing Square North | 40.751873 | -73.977706 | 5.5 |
| 1 | 20 | Pershing Square North | 40.751873 | -73.977706 | 2.75 |
| 1 | 21 | Pershing Square North | 40.751873 | -73.977706 | 3.5 |
| 1 | 22 | Pershing Square North | 40.751873 | -73.977706 | 2.25 |
| 1 | 23 | Pershing Square North | 40.751873 | -73.977706 | 3.0 |
+-----+-----+-----+-----+-----+-----+
23 rows selected (32.929 seconds)
0: jdbc:hive2://babar.es.its.nyu.edu:10000/>

```

5. Track the volume changing trends on a stable location on the mean of all weekdays

```

FROM average1 SELECT address,lon,lat,timeperiod,(sum(avgvolume)/5.0) AS meanAvg
WHERE ((weekname != "6" and weekname != "7") and address = "Pershing Square
North") GROUP BY address,lon,lat,timeperiod ORDER by cast(timeperiod as int)
asc

```

```

INFO : 2019-11-30 11:39:31.519 Stage-2 map = 0%, reduce = 0%
INFO : 2019-11-30 11:39:37.671 Stage-2 map = 100%, reduce = 0%, Cumulative CPU 1.56 sec
INFO : 2019-11-30 11:39:43.830 Stage-2 map = 100%, reduce = 100%, Cumulative CPU 4.68 sec
INFO : MapReduce Total cumulative CPU time: 4 seconds 680 msec
INFO : Ended Job = job_1569350662793_69747
INFO : MapReduce Jobs Launched:
INFO : Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 7.39 sec HDFS Read: 5494399 HDFS Write: 1632 SUCCESS
INFO : Stage-Stage-2: Map: 1 Reduce: 1 Cumulative CPU: 4.68 sec HDFS Read: 7332 HDFS Write: 1359 SUCCESS
INFO : Total MapReduce CPU Time Spent: 12 seconds 70 msec
INFO : Completed executing command(queryId=hive_20191130113838_de8403a0-bf63-497b-9739-7f7f71c64344): Time taken: 62.512 seconds
INFO : OK
+-----+-----+-----+-----+-----+
| address | lon | lat | timeperiod | meanavg |
+-----+-----+-----+-----+-----+
| Pershing Square North | 40.751873 | -73.977706 | 0 | 2.3033333333333333 |
| Pershing Square North | 40.751873 | -73.977706 | 1 | 1.8833333333333335 |
| Pershing Square North | 40.751873 | -73.977706 | 2 | 0.6 |
| Pershing Square North | 40.751873 | -73.977706 | 3 | 0.2 |
| Pershing Square North | 40.751873 | -73.977706 | 4 | 1.1733333333333333 |
| Pershing Square North | 40.751873 | -73.977706 | 5 | 4.77 |
| Pershing Square North | 40.751873 | -73.977706 | 6 | 15.63 |
| Pershing Square North | 40.751873 | -73.977706 | 7 | 41.71 |
| Pershing Square North | 40.751873 | -73.977706 | 8 | 79.62 |
| Pershing Square North | 40.751873 | -73.977706 | 9 | 74.37 |
| Pershing Square North | 40.751873 | -73.977706 | 10 | 24.23 |
| Pershing Square North | 40.751873 | -73.977706 | 11 | 15.37 |
| Pershing Square North | 40.751873 | -73.977706 | 12 | 15.09 |
| Pershing Square North | 40.751873 | -73.977706 | 13 | 17.318000000000002 |
| Pershing Square North | 40.751873 | -73.977706 | 14 | 16.169999999999998 |
| Pershing Square North | 40.751873 | -73.977706 | 15 | 22.169999999999998 |
| Pershing Square North | 40.751873 | -73.977706 | 16 | 30.669999999999998 |
| Pershing Square North | 40.751873 | -73.977706 | 17 | 64.79 |
| Pershing Square North | 40.751873 | -73.977706 | 18 | 60.510000000000005 |
| Pershing Square North | 40.751873 | -73.977706 | 19 | 29.419999999999998 |
| Pershing Square North | 40.751873 | -73.977706 | 20 | 15.45 |
| Pershing Square North | 40.751873 | -73.977706 | 21 | 8.370000000000001 |
| Pershing Square North | 40.751873 | -73.977706 | 22 | 6.1 |
| Pershing Square North | 40.751873 | -73.977706 | 23 | 4.17 |
+-----+-----+-----+-----+-----+
24 rows selected (63.507 seconds)
0: jdbc:hive2://babar.es.its.nyu.edu:10000/>

```