BIG DATA TECHNOLOGIES LAB

(ACADEMIC YEAR : 2017-2018)

I SEMESTER

ASSIGNMENT 6

TOPIC: APACHE HIVE - DML

ASHISH CHANDRAKANT DUSANE

M. TECH. (ACDS)

COMPUTER ENGG. DEPARTMENT

{ PRN : 170101261004 }

~: Topic: Apache Hive – DML :~

1. Create database company̓ located at /user/hive/mywarehouse/.

2. Create a hive managed table partitioned table employeė in company̓ database which can store the following data:

1,Anne,Admin,50000,A

2,Gokul,Admin,50000,B

3,Janet,Sales,60000,A

4,Hari,Admin,50000,C

5,Sanker,Admin,50000,C

6,Margaret,Tech,12000,A

7,Nirmal,Tech,12000,B

8,jinju,Engineer,45000,B

9,Nancy,Admin,50000,A

10,Andrew,Manager,40000,A

11,Arun,Manager,40000,B

12,Harish,Sales,60000,B

13,Robert,Manager,40000,A

14,Laura,Engineer,45000,A

15,Anju,Ceo,100000,B

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
hive>  CREATE  TABLE emp(emp_ID int, emp_Name string,Designation varchar(20),Salary f
loat,Grade varchar(10)) row format delimited fields terminated by ',';
OK
Time taken: 0.084 seconds
hive> LOAD DATA LOCAL INPATH '/home/student/Documents/emp_details.txt' into table emp
;
Loading data to table company.emp
Table company.emp stats: [numFiles=1, totalSize=349]
OK
Time taken: 0.175 seconds
hive> select * from emp;
OK
1       Anne    Admin   50000.0 A
2       Gokul   Admin   50000.0 B
3       Janet   Sales   60000.0 A
4       Hari    Admin   50000.0 C
5       Sanker  Admin   50000.0 C
6       Margaret        Tech    12000.0 A
7       Nirmal  Tech    12000.0 B
8       jinju   Engineer        45000.0 B
9       Nancy   Admin   50000.0 A
10      Andrew  Manager 40000.0 A
11      Arun    Manager 40000.0 B
12      Harish  Sales   60000.0 B
13      Robert  Manager 40000.0 A
14      Laura   Engineer        45000.0 A
15      Anju    Ceo     100000.0        B
Time taken: 0.122 seconds, Fetched: 15 row(s)
hive>
```

Partition the table using the fourth column.

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
hive> create table state_part(emp_ID int, emp_Name string,Designation varchar(20
),Grade varchar(10)) PARTITIONED BY (Salary float);
OK
Time taken: 0.084 seconds
hive> set hive.exec.dynamic.partition = TRUE;
hive> set hive.exec.dynamic.partition.mode=nonstrict;
hive> insert overwrite table state_part PARTITION(Salary) select emp_ID,emp_Name
,Designation,Grade,Salary from emp;
Query ID = student_20171223104246_332174d2-76c7-4741-a3f0-fbe0dc247d92
Total jobs = 3
```

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
         Time taken for load dynamic partitions : 895
        Loading partition {salary=50000.0}
        Loading partition {salary=45000.0}
        Loading partition {salary=40000.0}
        Loading partition {salary=60000.0}
        Loading partition {salary=100000.0}
        Loading partition {salary=12000.0}
         Time taken for adding to write entity : 32
Partition company.state_part{salary=100000.0} stats: [numFiles=1, numRows=2, tot
alSize=31, rawDataSize=29]
Partition company.state_part{salary=12000.0} stats: [numFiles=1, numRows=2, tota
lSize=34, rawDataSize=32]
Partition company.state_part{salary=40000.0} stats: [numFiles=1, numRows=4, tota
lSize=79, rawDataSize=75]
Partition company.state_part{salary=45000.0} stats: [numFiles=1, numRows=4, tota
lSize=83, rawDataSize=79]
Partition company.state_part{salary=50000.0} stats: [numFiles=1, numRows=6, tota
lSize=97, rawDataSize=91]
Partition company.state_part{salary=60000.0} stats: [numFiles=1, numRows=2, tota
lSize=34, rawDataSize=32]
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1   Cumulative CPU: 2.13 sec   HDFS Read: 4781 HDFS Write: 7
08 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 130 msec
OK
Time taken: 19.541 seconds
hive>
```

3. Save the above data in a text file on local file system and HDFS. Load data into employeė table from both LOCAL & HDFS filesystem.

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
hive>  CREATE  TABLE emp(emp_ID int, emp_Name string,Designation varchar(20),Salary f
loat,Grade varchar(10)) row format delimited fields terminated by ',';
OK
Time taken: 0.084 seconds
hive> LOAD DATA LOCAL INPATH '/home/student/Documents/emp_details.txt' into table emp
;
Loading data to table company.emp
Table company.emp stats: [numFiles=1, totalSize=349]
OK
Time taken: 0.175 seconds
hive> select * from emp;
OK
1       Anne    Admin   50000.0 A
2       Gokul   Admin   50000.0 B
3       Janet   Sales   60000.0 A
4       Hari    Admin   50000.0 C
5       Sanker  Admin   50000.0 C
6       Margaret        Tech    12000.0 A
7       Nirmal  Tech    12000.0 B
8       jinju   Engineer        45000.0 B
9       Nancy   Admin   50000.0 A
10      Andrew  Manager 40000.0 A
11      Arun    Manager 40000.0 B
12      Harish  Sales   60000.0 B
13      Robert  Manager 40000.0 A
14      Laura   Engineer        45000.0 A
15      Anju    Ceo     100000.0        B
Time taken: 0.122 seconds, Fetched: 15 row(s)
hive>
```

4. Add the following records to the text file and load the modified data into employee table using OVERWRITE.

LOAD DATA LOCAL INPATH "1987.csv" OVERWRITE INTO TABLE stg_airline.onTimePerf;

16,Aarathi,Manager,40000,B

17,Parvathy,Engineer,45000,B

18,Gopika,Admin,50000,B

19,Steven,Engineer,45000,A

20,Michael,Ceo,100000,A

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
hive> LOAD DATA LOCAL INPATH '/home/student/Documents/emp_details.txt'OVERWRITE
into table emp;
Loading data to table company.emp
Table company.emp stats: [numFiles=1, numRows=0, totalSize=480, rawDataSize=0]
OK
Time taken: 1.718 seconds
hive> SELECT * FROM EMP;
OK
1       Anne    Admin   50000.0 A
2       Gokul   Admin   50000.0 B
3       Janet   Sales   60000.0 A
4       Hari    Admin   50000.0 C
5       Sanker  Admin   50000.0 C
6       Margaret        Tech    12000.0 A
7       Nirmal  Tech    12000.0 B
8       jinju   Engineer        45000.0 B
9       Nancy   Admin   50000.0 A
10      Andrew  Manager 40000.0 A
11      Arun    Manager 40000.0 B
12      Harish  Sales   60000.0 B
13      Robert  Manager 40000.0 A
14      Laura   Engineer        45000.0 A
15      Anju    Ceo     100000.0        B
16      Aarathi Manager 40000.0 B
17      Parvathy        Engineer        45000.0 B
18      Gopika  Admin   50000.0 B
19      Steven  Engineer        45000.0 A
20      Michael Ceo     100000.0        A
Time taken: 0.448 seconds, Fetched: 20 row(s)
hive>
```

5. Create another table new_employees with the following records:

12,Priyanka,Admin,40000,C

22,Paras,Engineer,45000,B

23,Gopal,Sales,50000,C

24,Sukant,Engineer,45000,A

25,Murugan,CFO,100000,A

Append these records into employees table using INSERT.

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
hive> insert into table new_emp (emp_id, emp_name, designation, salary, gra
lues (25, 'Murugan', 'CFO', 100000, 'A');
Query ID = student_20171223110040_db24518b-c0be-47ba-8135-a6c338bb504a
Total jobs = 3
Launching Job 1 out of 3
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1514003803554_0006, Tracking URL = http://student-OptiPl
20:8088/proxy/application_1514003803554_0006/
Kill Command = /home/student/Documents/hadoop-2.7.3/bin/hadoop job  -kill j
14003803554_0006
```

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
hive> select * from new_emp;
OK
12      Priyanka        Admin   40000.0 C
22      Paras   Engineer        45000.0 B
23      Gopal   Sales   50000.0 C
24      Sukant  Engineer        45000.0 A
25      Murugan CFO     100000.0        A
Time taken: 0.08 seconds, Fetched: 5 row(s)
hive>
```

6. List the employees having salary>50000.

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
hive> select * from emp where Salary>50000;
OK
3       Janet   Sales   60000.0 A
12      Harish  Sales   60000.0 B
15      Anju    Ceo     100000.0        B
20      Michael Ceo     100000.0        A
Time taken: 0.09 seconds, Fetched: 4 row(s)
hive>
```

7. Select the list of employees whose names start from A.

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
hive> select * from emp where emp_Name like 'A%';
OK
1       Anne    Admin   50000.0 A
10      Andrew  Manager 40000.0 A
11      Arun    Manager 40000.0 B
15      Anju    Ceo     100000.0        B
16      Aarathi Manager 40000.0 B
Time taken: 0.1 seconds, Fetched: 5 row(s)
hive>
```

8. List the number of employees for each Designation.

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
Time taken: 0.1 seconds, Fetched: 5 row(s)
hive> select Designation,count(*) from emp group by Designation;
Query ID = student_20171223111131_7e83647c-fa86-4141-98ac-b0fc6ef6ed83
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks not specified. Estimated from input data size: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1514003803554_0007, Tracking URL = http://student-OptiPl
ex-3020:8088/proxy/application_1514003803554_0007/
```

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 4.12 sec   HDFS Read: 81
92 HDFS Write: 50 SUCCESS
Total MapReduce CPU Time Spent: 4 seconds 120 msec
OK
Admin   6
Ceo     2
Engineer        4
Manager 4
Sales   2
Tech    2
Time taken: 18.569 seconds, Fetched: 6 row(s)
hive>
```

9. Order the list of employees according to their names.

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
led rule: 'identifier' in table or column identifier
hive> select * from emp order by emp_Name;
Query ID = student_20171223111948_70926390-e674-4239-842c-6ade4591077a
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
```

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 3.6 sec   HDFS Read: 793
2 HDFS Write: 520 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 600 msec
OK
16      Aarathi Manager 40000.0 B
10      Andrew  Manager 40000.0 A
15      Anju    Ceo     100000.0        B
1       Anne    Admin   50000.0 A
11      Arun    Manager 40000.0 B
2       Gokul   Admin   50000.0 B
18      Gopika  Admin   50000.0 B
4       Hari    Admin   50000.0 C
12      Harish  Sales   60000.0 B
3       Janet   Sales   60000.0 A
14      Laura   Engineer        45000.0 A
6       Margaret        Tech    12000.0 A
20      Michael Ceo     100000.0        A
9       Nancy   Admin   50000.0 A
7       Nirmal  Tech    12000.0 B
17      Parvathy        Engineer        45000.0 B
13      Robert  Manager 40000.0 A
5       Sanker  Admin   50000.0 C
19      Steven  Engineer        45000.0 A
8       jinju   Engineer        45000.0 B
Time taken: 17.758 seconds, Fetched: 20 row(s)
hive>
```

10. Partition the employees table based upon the salary.

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
hive> create table state_part(emp_ID int, emp_Name string,Designation varchar(20
),Grade varchar(10)) PARTITIONED BY (Salary float);
OK
Time taken: 0.084 seconds
hive> set hive.exec.dynamic.partition = TRUE;
hive> set hive.exec.dynamic.partition.mode=nonstrict;
hive> insert overwrite table state_part PARTITION(Salary) select emp_ID,emp_Name
,Designation,Grade,Salary from emp;
Query ID = student_20171223104246_332174d2-76c7-4741-a3f0-fbe0dc247d92
Total jobs = 3
```

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
         Time taken for load dynamic partitions : 895
        Loading partition {salary=50000.0}
        Loading partition {salary=45000.0}
        Loading partition {salary=40000.0}
        Loading partition {salary=60000.0}
        Loading partition {salary=100000.0}
        Loading partition {salary=12000.0}
         Time taken for adding to write entity : 32
Partition company.state_part{salary=100000.0} stats: [numFiles=1, numRows=2, tot
alSize=31, rawDataSize=29]
Partition company.state_part{salary=12000.0} stats: [numFiles=1, numRows=2, tota
lSize=34, rawDataSize=32]
Partition company.state_part{salary=40000.0} stats: [numFiles=1, numRows=4, tota
lSize=79, rawDataSize=75]
Partition company.state_part{salary=45000.0} stats: [numFiles=1, numRows=4, tota
lSize=83, rawDataSize=79]
Partition company.state_part{salary=50000.0} stats: [numFiles=1, numRows=6, tota
lSize=97, rawDataSize=91]
Partition company.state_part{salary=60000.0} stats: [numFiles=1, numRows=2, tota
lSize=34, rawDataSize=32]
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1   Cumulative CPU: 2.13 sec   HDFS Read: 4781 HDFS Write: 7
08 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 130 msec
OK
Time taken: 19.541 seconds
hive>
```

11. List to top 5 highly paid employees.

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
hive> select max(salary) from emp limit 5;
Query ID = student_20171223113041_36053a5f-9ab7-429b-bf64-85878350f9ae
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
```

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 3.61 sec   HDFS Read: 80
27 HDFS Write: 122 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 610 msec
OK
15      Anju    Ceo     100000.0        B
20      Michael Ceo     100000.0        A
12      Harish  Sales   60000.0 B
3       Janet   Sales   60000.0 A
1       Anne    Admin   50000.0 A
Time taken: 18.361 seconds, Fetched: 5 row(s)
hive>
```

12. List all admins whose salary is > 45000.

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
hive> select * from emp where designation="Admin" and salary>45000;
OK
1       Anne    Admin   50000.0 A
2       Gokul   Admin   50000.0 B
4       Hari    Admin   50000.0 C
5       Sanker  Admin   50000.0 C
9       Nancy   Admin   50000.0 A
18      Gopika  Admin   50000.0 B
Time taken: 0.084 seconds, Fetched: 6 row(s)
hive>
```

13. Compute the average salary paid by the company to its employees.

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
Time taken: 18.841 seconds, Fetched: 1 row(s)
hive> select avg(salary) from emp;
Query ID = student_20171223115254_ce7a3da0-5a4e-4890-8eb9-339632dd2367
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 1
```

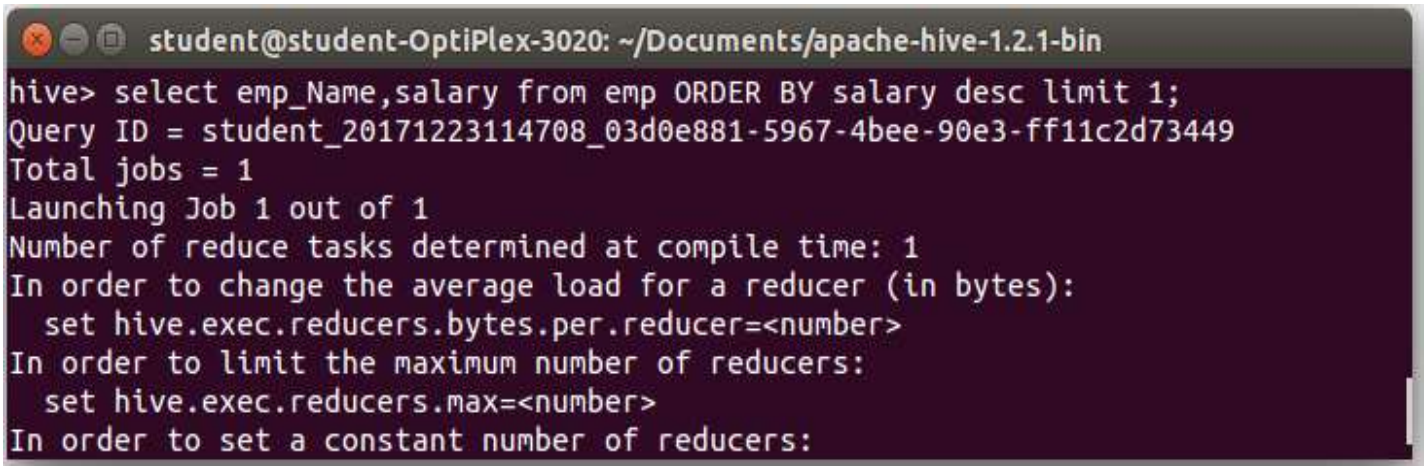
```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
S Write: 8 SUCCESS
Total MapReduce CPU Time Spent: 4 seconds 90 msec
OK
49200.0
Time taken: 19.038 seconds, Fetched: 1 row(s)
hive>
```

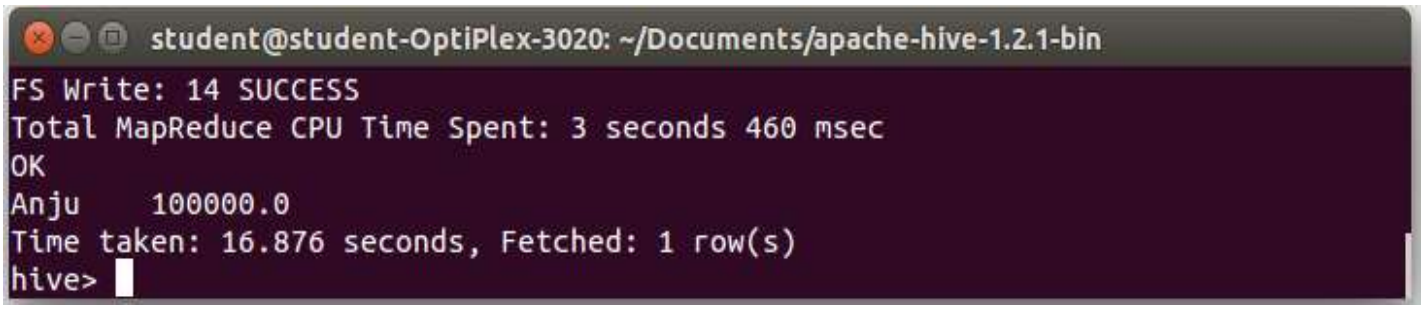
14. Compute the total salary paid by the company per month.

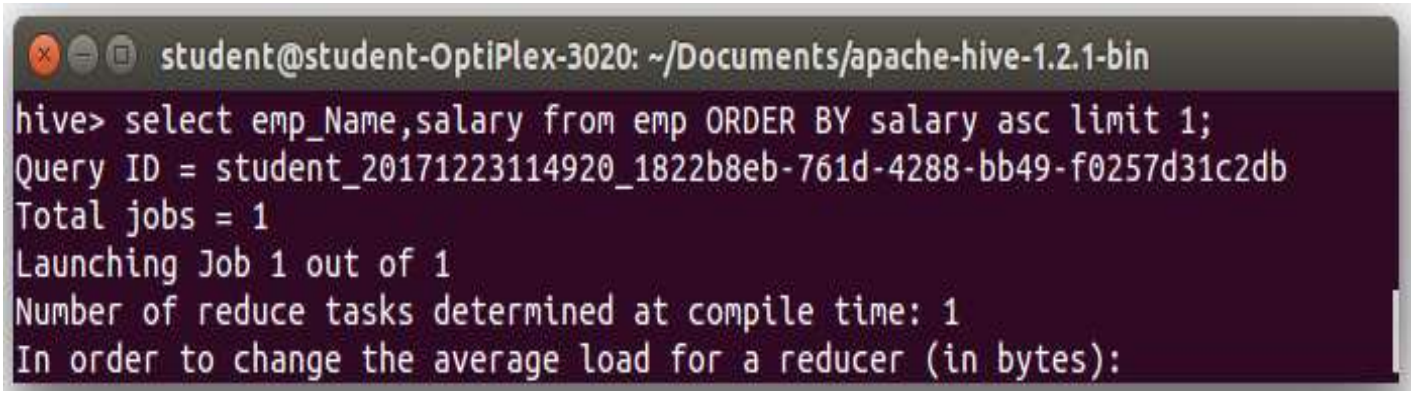
```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
hive> select sum(salary) from emp;
Query ID = student_20171223120407_e312ec67-4869-4512-9c43-223cf5e2e0c0
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
```

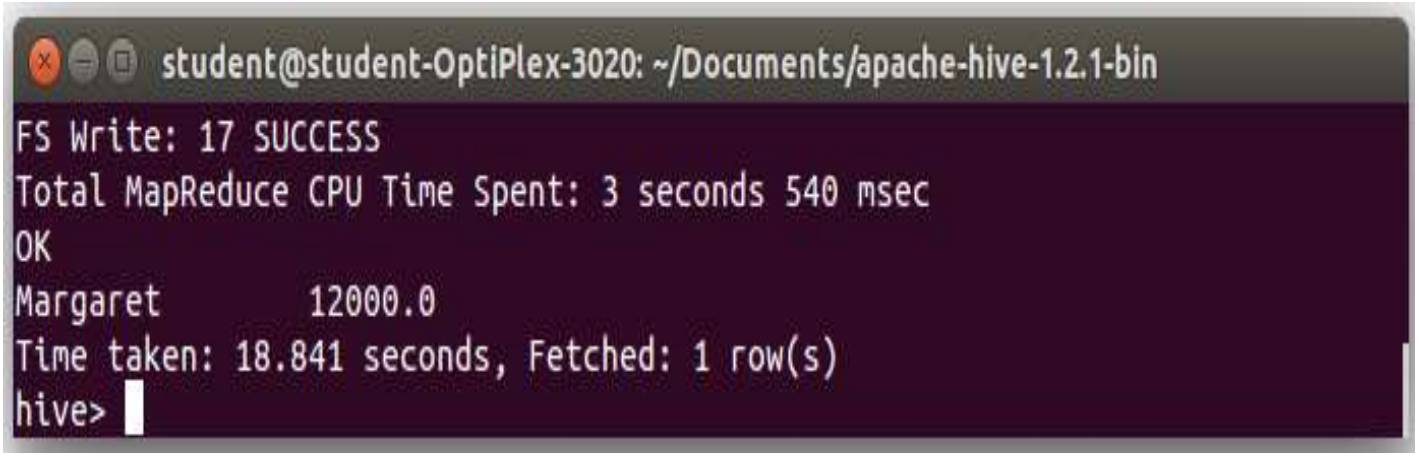
```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 4.14 sec   HDFS Read: 7927 HDF
S Write: 9 SUCCESS
Total MapReduce CPU Time Spent: 4 seconds 140 msec
OK
984000.0
Time taken: 19.06 seconds, Fetched: 1 row(s)
hive>
```

15. List the employee with the highest salary. List the employee with the minimum salary.









16. Display DISTINCT salaries paid by the company.

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
S Write: 49 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 750 msec
OK
12000.0
40000.0
45000.0
50000.0
60000.0
100000.0
Time taken: 18.349 seconds, Fetched: 6 row(s)
hive> select DISTINCT(salary) from emp;
```

17. List the employees with increasing order of salary paid.

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
Query ID = student_20171223120042_07e635b9-eef3-42bb-b264-7932f64df6c4
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
```

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
S Write: 295 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 230 msec
OK
Margaret        12000.0
Nirmal  12000.0
Andrew  40000.0
Aarathi 40000.0
Robert  40000.0
Arun    40000.0
Steven  45000.0
jinju   45000.0
Parvathy        45000.0
Laura   45000.0
Anne    50000.0
Gopika  50000.0
Nancy   50000.0
Sanker  50000.0
Hari    50000.0
Gokul   50000.0
Harish  60000.0
Janet   60000.0
Michael 100000.0
Anju    100000.0
Time taken: 17.295 seconds, Fetched: 20 row(s)
hive>
```

18. Make two partitions of the table with the following criteria:

A) Partition 1 consisting of employees having salary <=50000

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
Time taken: 0.953 seconds
hive> create table state_part1(emp_ID int, emp_Name string,Designation varchar(20
),Grade varchar(10)) PARTITIONED BY (Salary float);
OK
Time taken: 0.121 seconds
hive> set hive.exec.dynamic.partition = TRUE;
hive> set hive.exec.dynamic.partition.mode=nonstrict;
hive> insert overwrite table state_part1 PARTITION(Salary) select emp_ID,emp_Name
,Designation,Grade,Salary from emp where Salary<=50000;
```

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
         Time taken for load dynamic partitions : 433
        Loading partition {salary=12000.0}
        Loading partition {salary=40000.0}
        Loading partition {salary=45000.0}
        Loading partition {salary=50000.0}
         Time taken for adding to write entity : 0
Partition company.state_part1{salary=12000.0} stats: [numFiles=1, numRows=2, tota
lSize=34, rawDataSize=32]
Partition company.state_part1{salary=40000.0} stats: [numFiles=1, numRows=4, tota
lSize=79, rawDataSize=75]
Partition company.state_part1{salary=45000.0} stats: [numFiles=1, numRows=4, tota
lSize=83, rawDataSize=79]
Partition company.state_part1{salary=50000.0} stats: [numFiles=1, numRows=6, tota
lSize=97, rawDataSize=91]
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1   Cumulative CPU: 3.06 sec   HDFS Read: 5315 HDFS Write: 54
2 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 60 msec
OK
Time taken: 14.948 seconds
hive>
```

B) Partition 2 consisting of employees having salary > 50000

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
Time taken: 0.953 seconds
hive> create table state_part1(emp_ID int, emp_Name string,Designation varchar(20
),Grade varchar(10)) PARTITIONED BY (Salary float);
OK
Time taken: 0.121 seconds
hive> set hive.exec.dynamic.partition = TRUE;
hive> set hive.exec.dynamic.partition.mode=nonstrict;
hive> insert overwrite table state_part1 PARTITION(Salary) select emp_ID,emp_Name
,Designation,Grade,Salary from emp where Salary<=50000;
```

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
Moving data to: hdfs://localhost:9000/user/hive/warehouse/company.db/state_part1/
.hive-staging_hive_2017-12-23_14-18-52_951_4342634158954283683-1/-ext-10000
Loading data to table company.state_part1 partition (salary=null)
         Time taken for load dynamic partitions : 244
        Loading partition {salary=60000.0}
        Loading partition {salary=100000.0}
         Time taken for adding to write entity : 0
Partition company.state_part1{salary=100000.0} stats: [numFiles=1, numRows=2, tot
alSize=31, rawDataSize=29]
Partition company.state_part1{salary=60000.0} stats: [numFiles=1, numRows=2, tota
lSize=34, rawDataSize=32]
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1   Cumulative CPU: 3.12 sec   HDFS Read: 5383 HDFS Write: 20
9 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 120 msec
OK
Time taken: 13.198 seconds
hive> insert overwrite table state_part1 PARTITION(Salary) select emp_ID,emp_Name
,Designation,Grade,Salary from emp where Salary>50000;
```

19. List the employees having salary between 50000 and 60000 in partition 2.

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
hive> select salary,emp_name from state_part1 where salary BETWEEN 50000 AND 6000
0;
OK
50000.0 Anne
50000.0 Gokul
50000.0 Hari
50000.0 Sanker
50000.0 Nancy
50000.0 Gopika
60000.0 Janet
60000.0 Harish
Time taken: 0.111 seconds, Fetched: 8 row(s)
hive>
```

20. Create partition 3 for all Engineer having salary >=45000.

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
hive> create table state_part2(emp_ID int, emp_Name string,Grade varchar(10)) PART
ITIONED BY (Salary float,Designation varchar(20));
OK
Time taken: 0.142 seconds
hive> set hive.exec.dynamic.partition = TRUE;
hive> set hive.exec.dynamic.partition.mode=nonstrict;
hive> insert overwrite table state_part PARTITION(Salary,Designation) select emp_I
```

```
student@student-OptiPlex-3020: ~/Documents/apache-hive-1.2.1-bin
hive-staging_hive_2017-12-23_14-31-03_607_3734833133284335793-1/-ext-10000
Loading data to table company.state_part2 partition (salary=null, designation=null
)
         Time taken for load dynamic partitions : 145
        Loading partition {salary=45000.0, designation=Engineer}
         Time taken for adding to write entity : 0
Partition company.state_part2{salary=45000.0, designation=Engineer} stats: [numFil
es=1, numRows=4, totalSize=47, rawDataSize=43]
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1   Cumulative CPU: 3.31 sec   HDFS Read: 5479 HDFS Write: 158
 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 310 msec
OK
Time taken: 12.786 seconds
hive> insert overwrite table state_part2 PARTITION(Salary,Designation) select emp_
ID,emp_Name,Grade,Salary,Designation from emp where Salary>=45000 AND Designation=
"Engineer";
```