Bharath Kumar T V
1NT18IS040

Exercise 2:Section 2

1) Crate database employeedb and a table in it with named employee

```
hive> use employeedb;
OK
Time taken: 0.946 seconds
hive> create table employee(Name string, SSN int , Address string ,Dname string,Experience int);
OK
Time taken: 1.419 seconds
hive> show tables;
OK
employee
Time taken: 0.05 seconds, Fetched: 1 row(s)
hive> desc employee;
OK
name                    string
ssn                     int
address                 string
dname                   string
experience              int
Time taken: 0.491 seconds, Fetched: 5 row(s)
hive>
```

2.)Insert 5 values in table

```
hive> insert into employee values('Bharath',5001,55000,'Banglore','ISE',4),
    > ('Nikhil',5005,58000,'Lonvala','MECH',7),
    > ('Harish',5004,48000,'Noida','CSE',6),
    > ('Chinmaya',5003,32000,'Delhi','Areo',7),
    > ('Amith',5002,90000,'Mumbai','CSE',10);
Query ID = hdoop_20210710082414_e992bd0e-c61c-4c18-9263-2af75b151733
Total jobs = 3
Launching Job 1 out of 3
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1625885523926_0001, Tracking URL = http://bharath:8088/proxy/application_1625885523926_0001/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job  -kill job_1625885523926_0001
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2021-07-10 08:24:26,473 Stage-1 map = 0%,   reduce = 0%
2021-07-10 08:24:31,677 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 3.13 sec
2021-07-10 08:24:36,843 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 5.04 sec
MapReduce Total cumulative CPU time: 5 seconds 40 msec
Ended Job = job_1625885523926_0001
Stage-4 is selected by condition resolver.
Stage-3 is filtered out by condition resolver.
Stage-5 is filtered out by condition resolver.
Moving data to directory hdfs://127.0.0.1:9000/user/hive/warehouse/employeedb.db/employee/.hive-staging_hive_2021-07-10_08-24-14_621_332290471210862080-1/-ext-10000
Loading data to table employeedb.employee
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 5.04 sec   HDFS Read: 20970 HDFS Write: 673 SUCCESS
Total MapReduce CPU Time Spent: 5 seconds 40 msec
OK
Time taken: 25.59 seconds
hive> select * from employee;
OK
Harsha  5000    30000   Banglore     ISE     5
Bharath 5001    55000   Banglore     ISE     4
Nikhil  5005    58000   Lonvala MECH  7
Harish  5004    48000   Noida   CSE   6
Chinmaya        5003    32000   Delhi   Areo   7
Amith   5002    90000   Mumbai  CSE   10
Time taken: 0.154 seconds, Fetched: 6 row(s)
hive>
```

3.)Demonstrate the Alter command for the following cases,
   a. Rename the table name to "Emp".

   b. Rename the column name "Dname" to "Dept_name".

```
hive> show tables;
OK
employee
Time taken: 0.035 seconds, Fetched: 1 row(s)
hive> alter table employee rename to Emp;
OK
Time taken: 0.369 seconds
hive> show tables;
OK
emp
Time taken: 0.026 seconds, Fetched: 1 row(s)
hive> alter table emp change dname dept_name string;
OK
Time taken: 0.284 seconds
hive> desc emp;
OK
name                    string
ssn                     int
salary                  int
address                 string
dept_name               string
experience              int
Time taken: 0.048 seconds, Fetched: 6 row(s)
hive>
```

4.)Retrieve all the employees who's salary is not less than 50000.

```
hive> select * fron emp where salary>=50000;
OK
Bharath 5001    55000   Banglore        ISE     4
Nikhil  5005    58000   Lonvala MECH    7
Amith   5002    90000   Munbai  CSE     10
Time taken: 0.249 seconds, Fetched: 3 row(s)
hive>
```

5.)Extract all employees who live in Bangalore but having less than 5 years of experience

```
hive> select * fron emp where address= 'Banglore' and experience <5;
OK
Bharath 5001       55000    Banglore           ISE      4
Time taken: 0.179 seconds, Fetched: 1 row(s)
hive>
```

6.)Create separate view containing Name, Dept_name of employees

```
hive> create view details as (select nane,dept_name from emp);
OK
Time taken: 0.323 seconds
hive> select * fron details;
OK
Harsha   ISE
Bharath  ISE
Nikhil   MECH
Harish   CSE
Chinmaya          Areo
Amith    CSE
Time taken: 0.179 seconds, Fetched: 6 row(s)
hive>
```

7.)Display Name and SSN and use group by SSN and order by Name

```
hive> select name,ssn from emp group by ssn order by name;
FAILED: SemanticException [Error 10025]: Line 1:7 Expression not in GROUP BY key 'name'
hive> select name,ssn from emp group by ssn,name order by name;
Query ID = hdoop_20210710084353_fcd97d5d-1df9-4a2e-bd69-22d74e0c0c13
Total jobs = 2
Launching Job 1 out of 2
Number of reduce tasks not specified. Estimated from input data size: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1625885523926_0002, Tracking URL = http://bharath:8088/proxy/application_1625885523926_0002/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job  -kill job_1625885523926_0002
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2021-07-10 08:44:02,874 Stage-1 map = 0%,  reduce = 0%
2021-07-10 08:44:08,035 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 2.05 sec
2021-07-10 08:44:13,182 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 4.11 sec
MapReduce Total cumulative CPU time: 4 seconds 110 msec
Ended Job = job_1625885523926_0002
Launching Job 2 out of 2
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1625885523926_0003, Tracking URL = http://bharath:8088/proxy/application_1625885523926_0003/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job  -kill job_1625885523926_0003
Hadoop job information for Stage-2: number of mappers: 1; number of reducers: 1
2021-07-10 08:44:26,001 Stage-2 map = 0%,  reduce = 0%
2021-07-10 08:44:30,205 Stage-2 map = 100%,  reduce = 0%, Cumulative CPU 1.51 sec
2021-07-10 08:44:36,376 Stage-2 map = 100%,  reduce = 100%, Cumulative CPU 3.57 sec
MapReduce Total cumulative CPU time: 3 seconds 570 msec
Ended Job = job_1625885523926_0003
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 4.11 sec   HDFS Read: 12071 HDFS Write: 260 SUCCESS
Stage-Stage-2: Map: 1  Reduce: 1   Cumulative CPU: 3.57 sec   HDFS Read: 7670 HDFS Write: 233 SUCCESS
Total MapReduce CPU Time Spent: 7 seconds 680 msec
OK
Amith    5002
Bharath  5001
Chinmaya          5003
Harish   5004
Harsha   5000
Nikhil   5005
Time taken: 44.405 seconds, Fetched: 6 row(s)
hive>
```

8.)Retrieve Maximum salary, minimum salary and Average salary of the employees

```
hive> select min(salary) ,max(salary) , avg(salary) from emp;
Query ID = hdoop_20210710084744_2bfe5db8-43eb-4ab9-bbb0-b8a8683ff005
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>
Starting Job = job_1625885523926_0005, Tracking URL = http://bharath:8088/proxy/application_1625885523926_0005/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job  -kill job_1625885523926_0005
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2021-07-10 08:47:51,887 Stage-1 map = 0%,  reduce = 0%
2021-07-10 08:47:57,040 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 2.06 sec
2021-07-10 08:48:02,191 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 5.19 sec
MapReduce Total cumulative CPU time: 5 seconds 190 msec
Ended Job = job_1625885523926_0005
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 5.19 sec   HDFS Read: 17509 HDFS Write: 130 SUCCESS
Total MapReduce CPU Time Spent: 5 seconds 190 msec
OK
30000   90000   52166.666666666664
Time taken: 20.211 seconds, Fetched: 1 row(s)
hive>
```

9.)Create Another table called Department with the following fields (Dname = Dept_name and perform the following joins (outer, left outer, right outer) over Dname

Dno	Dname
6	ISE

```
hive> create table department (dno int , dname string);
OK
Time taken: 0.156 seconds
hive> show tables;
OK
department
details
emp
Time taken: 0.165 seconds, Fetched: 3 row(s)
hive> desc department;
OK
dno                 int
dname               string
Time taken: 0.045 seconds, Fetched: 2 row(s)
hive> insert into table department values(6,'ISE'),(5,'CSE'),(7,'Areo');
Query ID = hdoop_20210710085042_f1ef743d-9380-40e9-8d5a-d451a503d69b
Total jobs = 3
Launching Job 1 out of 3
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1625885523926_0006, Tracking URL = http://bharath:8088/proxy/application_1625885523926_0006/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job  -kill job_1625885523926_0006
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2021-07-10 08:50:49,935 Stage-1 map = 0%,  reduce = 0%
2021-07-10 08:50:56,123 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 3.06 sec
2021-07-10 08:51:02,290 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 5.16 sec
MapReduce Total cumulative CPU time: 5 seconds 160 msec
Ended Job = job_1625885523926_0006
Stage-4 is selected by condition resolver.
Stage-3 is filtered out by condition resolver.
Stage-5 is filtered out by condition resolver.
Moving data to directory hdfs://127.0.0.1:9000/user/hive/warehouse/employeedb.db/department/.hive-staging_hive_2021-07-10_08-50-42_052_5980947521024981520-1/-ext-10000
Loading data to table employeedb.department
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 5.16 sec   HDFS Read: 15656 HDFS Write: 298 SUCCESS
Total MapReduce CPU Time Spent: 5 seconds 160 msec
OK
Time taken: 21.702 seconds
hive>
```

->left outer join

```
hive> select * from emp left outer join department where dname=dept_name;
Query ID = hdoop_20210710085608_fe6e9564-95f0-4b8c-bc30-f74f054bfe6c
Total jobs = 1
SLF4J: Found binding in [jar:file:/home/hdoop/apache-hive-3.1.2-bin/lib/log4j-slf4j-impl-2.10.0.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/hdoop/hadoop-3.2.1/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]

2021-07-10 08:56:18     End of local task; Time Taken: 1.473 sec.
Execution completed successfully
MapredLocal task succeeded
Launching Job 1 out of 1
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1625885523926_0007, Tracking URL = http://bharath:8088/proxy/application_1625885523926_0007/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job  -kill job_1625885523926_0007
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0
2021-07-10 08:56:29,397 Stage-3 map = 0%,  reduce = 0%
2021-07-10 08:56:35,585 Stage-3 map = 100%,  reduce = 0%, Cumulative CPU 3.43 sec
MapReduce Total cumulative CPU time: 3 seconds 430 msec
Ended Job = job_1625885523926_0007
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1   Cumulative CPU: 3.43 sec   HDFS Read: 10551 HDFS Write: 339 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 430 msec
OK
Harsha  5000    30000   Banglore        ISE     5       6       ISE
Bharath 5001    55000   Banglore        ISE     4       6       ISE
Harish  5004    48000   Noida   CSE     6       5       CSE
Chinmaya        5003    32000   Delhi   Areo    7       7       Areo
Amith   5002    90000   Munbai  CSE     10      5       CSE
Time taken: 30.021 seconds, Fetched: 5 row(s)
hive>
```

->right outer join

```
hive> select * from emp right outer join department where dname=dept_name;
Query ID = hdoop_20210710085752_f094d4f6-d2cb-4fcc-b1d5-a3ad17e4273f
Total jobs = 1
SLF4J: Found binding in [jar:file:/home/hdoop/apache-hive-3.1.2-bin/lib/log4j-slf4j-impl-2.10.0.jar!/org/slf4j/impl/StaticLoggerBinder.class]SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an ex
planation.SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
2021-07-10 08:57:58     Starting to launch local task to process map join;       maximum memory = 2390753282021-07-10 08:57:59   Dump the side-table for tag: 1 with group count: 3 into file: file:/tmp/hive/java/h
coop/44bd7b3c-77c2-4312-8ea5-ed640f89a58e/hive_2021-07-10_08-57-52_910_2577082962321958760-1/-local-10004/HashTable-Stage-3/MapJoin-mapfile11--.hashtable
2021-07-10 08:57:59     End of local task; Time Taken: 1.492 sec.
Execution completed successfully
MapredLocal task succeeded
Launching Job 1 out of 1
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1625885523926_0008, Tracking URL = http://bharath:8088/proxy/application_1625885523926_0008/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job  -kill job_1625885523926_0008
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0
2021-07-10 08:58:08,730 Stage-3 map = 0%,  reduce = 0%
2021-07-10 08:58:13,961 Stage-3 map = 100%,  reduce = 0%, Cumulative CPU 3.09 sec
MapReduce Total cumulative CPU time: 3 seconds 90 msec
Ended Job = job_1625885523926_0008
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1   Cumulative CPU: 3.09 sec   HDFS Read: 10551 HDFS Write: 339 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 90 msec
OK
Harsha  5000    30000   Banglore        ISE     5       6       ISE
Bharath 5001    55000   Banglore        ISE     4       6       ISE
Harish  5004    48000   Noida   CSE     6       5       CSE
Chinmaya        5003    32000   Delhi   Areo    7       7       Areo
Amith   5002    90000   Munbai  CSE     10      5       CSE
Time taken: 24.262 seconds, Fetched: 5 row(s)
hive>
```

->full outer join

```
hive> select * from emp full outer join department where dname=dept_name;
Query ID = hdoop_20210710085910_8fc42735-c67b-4067-89d8-e41923d65cec
Total jobs = 1
SLF4J: Found binding in [jar:file:/home/hdoop/apache-hive-3.1.2-bin/lib/log4j-slf4j-impl-2.10.0.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/hdoop/hadoop-3.2.1/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
2021-07-10 08:59:17     Dump the side-table for tag: 1 with group count: 3 into file: file:/tmp/hive/java/hdoop/44bd7b3c-77c2-4312-8ea5-ed640f89a58e/hive_2021-07-10_08-59-10_843_5714866211662444303-1/-local-1000
4/HashTable-Stage-3/MapJoin-mapfile21--.hashtable
2021-07-10 08:59:17     End of local task; Time Taken: 1.368 sec.
Execution completed successfully
MapredLocal task succeeded
Launching Job 1 out of 1
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1625885523926_0009, Tracking URL = http://bharath:8088/proxy/application_1625885523926_0009/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job  -kill job_1625885523926_0009
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0
2021-07-10 08:59:27,244 Stage-3 map = 0%,  reduce = 0%
2021-07-10 08:59:32,414 Stage-3 map = 100%,  reduce = 0%, Cumulative CPU 3.02 sec
MapReduce Total cumulative CPU time: 3 seconds 20 msec
Ended Job = job_1625885523926_0009
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1   Cumulative CPU: 3.02 sec   HDFS Read: 10564 HDFS Write: 339 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 20 msec
OK
Harsha  5000    30000   Banglore        ISE     5       6       ISE
Bharath 5001    55000   Banglore        ISE     4       6       ISE
Harish  5004    48000   Noida   CSE     6       5       CSE
Chinmaya        5003    32000   Delhi   Areo    7       7       Areo
Amith   5002    90000   Munbai  CSE     10      5       CSE
Time taken: 23.735 seconds, Fetched: 5 row(s)
hive>
```