Hive assignment

Bharath kumar T V

1.)Insert five records

```
hive> insert into table Bank values ( 22, "axis","manglore"),
    > (23 , "sbi" , "Delhi"),
    > (24 , "icici" , "Mumbai"),
    > (25 , "pkgb" , "Hyderbad");
Query ID = hdoop_20210705081001_ec600e7c-8474-4070-ac6e-93c44488654c
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_1625450345081_0002, Tracking URL = http://bharath:8088/proxy/application_1625450345081_0002/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job  -kill job_1625450345081_0002
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2021-07-05 08:10:08,510 Stage-1 map = 0%,  reduce = 0%
2021-07-05 08:10:13,669 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 3.18 sec
2021-07-05 08:10:18,818 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 5.05 sec
MapReduce Total cumulative CPU time: 5 seconds 50 msec
Ended Job = job_1625450345081_0002
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/organisation.db/bank/.hive-staging_hive_2021-07-05_08-10-01_529_49086287957900324-1/-ext-10000
Loading data to table organisation.bank
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 5.05 sec   HDFS Read: 17092 HDFS Write: 385 SUCCESS
Total MapReduce CPU Time Spent: 5 seconds 50 msec
OK
Time taken: 20.872 seconds
hive> select * from bank;
OK
20      hdfc    Banglore
22      axis    manglore
23      sbi     Delhi
24      icici   Mumbai
25      pkgb    Hyderbad
Time taken: 0.147 seconds, Fetched: 5 row(s)
```

2.)Demonstrate the Alter command for the above tables

```
hive> alter table Account rename to Accounts;
OK
Time taken: 0.128 seconds
hive> show tables;
OK
accounts
bank
customer
Time taken: 0.057 seconds, Fetched: 3 row(s)
hive> alter table bank change blocation location string;
OK
Time taken: 0.135 seconds
hive> desc bank;
OK
bank_id                 int
bname                   string
location                string
Time taken: 0.039 seconds, Fetched: 3 row(s)
hive> show tables;
OK
accounts
bank
customer
Time taken: 0.024 seconds, Fetched: 3 row(s)
hive>
```

3.)Retrieve all the customers who have account in AXIS bank

```
hive> select customer.cust_id , cname from customer,accounts,bank
    > where customer.accid = accounts.accid and accounts.bnk_id = bank.bank_id
    > and bname = 'axis';
No Stats for organisation@customer, Columns: cname, accid, cust_id
Query ID = hdoop_20210705134940_530bb21c-7c2b-4713-b2ce-1d87f268f98d
Total jobs = 1

SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]

2021-07-05 13:49:54     Dump the side-table for tag: 0 with group count: 2 into file: file:/tmp/hive/java/hdoop/30ec6acf-74c6-446d-a46b-d21e7cdfeec2/hive_2021-07-05_13-49-40_572_4736569057326564931-1/-local-10005/HashTable-Stage-5/MapJoin-mapfile10--.hashtable
2021-07-05 13:49:54     End of local task; Time Taken: 2.344 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_1625467450533_0006, Tracking URL = http://bharath:8088/proxy/application_1625467450533_0006/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job  -kill job_1625467450533_0006
Hadoop job information for Stage-5: number of mappers: 1; number of reducers: 0
2021-07-05 13:50:10,469 Stage-5 map = 0%,  reduce = 0%
2021-07-05 13:50:25,981 Stage-5 map = 100%,  reduce = 0%, Cumulative CPU 3.22 sec
MapReduce Total cumulative CPU time: 3 seconds 220 msec
Ended Job = job_1625467450533_0006
MapReduce Jobs Launched:
Stage-Stage-5: Map: 1   Cumulative CPU: 3.22 sec   HDFS Read: 12095 HDFS Write: 128 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 220 msec
OK
1       bharath
4       hari
Time taken: 54.567 seconds, Fetched: 2 row(s)
hive>
```

4.)Retrieve the customer with maximum income and minimum income

```
hive> select cust_id , cname from customer where income in(select min(income) from customer);
Query ID = hdoop_20210710110806_8e5e02e2-af64-4a54-9f18-81f0ea621ecb
Total jobs = 2
Launching Job 1 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_1625894617298_0001, Tracking URL = http://bharath:8088/proxy/application_1625894617298_0001/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job  -kill job_1625894617298_0001
Hadoop job information for Stage-2: number of mappers: 1; number of reducers: 1
2021-07-10 11:08:21,431 Stage-2 map = 0%,  reduce = 0%
2021-07-10 11:08:26,633 Stage-2 map = 100%,  reduce = 0%, Cumulative CPU 1.71 sec
2021-07-10 11:08:32,819 Stage-2 map = 100%,  reduce = 100%, Cumulative CPU 4.34 sec
MapReduce Total cumulative CPU time: 4 seconds 340 msec
Ended Job = job_1625894617298_0001
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 f
or an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
2021-07-10 11:08:41     Dump the side-table for tag: 0 with group count: 4 into file: file:/tmp/hive/java/hdoop/163edaf4-fd7e-4504-aeb1-299db3f44baf/hive_2021-07-10_11-08-06_266_5883743570237533608-1/-local-1000
5/HashTable-Stage-3/MapJoin-mapfile00--.hashtable
2021-07-10 11:08:41     End of local task; Time Taken: 1.327 sec.
Execution completed successfully
MapredLocal task succeeded
Launching Job 2 out of 2
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1625894617298_0002, Tracking URL = http://bharath:8088/proxy/application_1625894617298_0002/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job  -kill job_1625894617298_0002
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0
2021-07-10 11:08:50,734 Stage-3 map = 0%,  reduce = 0%
2021-07-10 11:08:54,852 Stage-3 map = 100%,  reduce = 0%, Cumulative CPU 2.13 sec
MapReduce Total cumulative CPU time: 2 seconds 130 msec
Ended Job = job_1625894617298_0002
MapReduce Jobs Launched:
Stage-Stage-2: Map: 1  Reduce: 1   Cumulative CPU: 4.34 sec   HDFS Read: 12929 HDFS Write: 117 SUCCESS
Stage-Stage-3: Map: 1   Cumulative CPU: 2.13 sec   HDFS Read: 7046 HDFS Write: 109 SUCCESS
Total MapReduce CPU Time Spent: 6 seconds 470 msec
OK
3       chinmay
Time taken: 50.755 seconds, Fetched: 1 row(s)
hive>
```

```
hive> select cust_id , cname,salary from customer where income in(select max(income) from customer);
FAILED: SemanticException [Error 10004]: Line 1:23 Invalid table alias or column reference 'salary': (possible column names are: cust_id, cname, income, accid, dob)
hive> select cust_id , cname,income from customer where income in(select max(income) from customer);
Query ID = hdoop_20210710110946_14052725-60c0-41ab-a64e-a27c1d517fbc
Total jobs = 2
Launching Job 1 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_1625894617298_0003, Tracking URL = http://bharath:8088/proxy/application_1625894617298_0003/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job  -kill job_1625894617298_0003
Hadoop job information for Stage-2: number of mappers: 1; number of reducers: 1
2021-07-10 11:09:55,907 Stage-2 map = 0%,  reduce = 0%
2021-07-10 11:10:00,080 Stage-2 map = 100%,  reduce = 0%, Cumulative CPU 1.86 sec
2021-07-10 11:10:06,271 Stage-2 map = 100%,  reduce = 100%, Cumulative CPU 4.8 sec
MapReduce Total cumulative CPU time: 4 seconds 800 msec
Ended Job = job_1625894617298_0003

SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
2021-07-10 11:10:15     Dump the side-table for tag: 0 with group count: 4 into file: file:/tmp/hive/java/hdoop/163edaf4-fd7e-4504-aeb1-299db3f44baf/hive_2021-07-10_11-09-46_139_1333128381021883289-1/-local-1000
5/HashTable-Stage-3/MapJoin-mapfile10--.hashtable
Execution completed successfully
MapredLocal task succeeded
Launching Job 2 out of 2
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1625894617298_0004, Tracking URL = http://bharath:8088/proxy/application_1625894617298_0004/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job  -kill job_1625894617298_0004
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0
2021-07-10 11:10:26,676 Stage-3 map = 0%,  reduce = 0%
2021-07-10 11:10:31,844 Stage-3 map = 100%,  reduce = 0%, Cumulative CPU 2.2 sec
MapReduce Total cumulative CPU time: 2 seconds 200 msec
Ended Job = job_1625894617298_0004
MapReduce Jobs Launched:
Stage-Stage-2: Map: 1  Reduce: 1   Cumulative CPU: 4.8 sec   HDFS Read: 12960 HDFS Write: 117 SUCCESS
Stage-Stage-3: Map: 1   Cumulative CPU: 2.2 sec   HDFS Read: 7079 HDFS Write: 117 SUCCESS
Total MapReduce CPU Time Spent: 7 seconds 0 msec
OK
1       bharath 34000.0
Time taken: 47.907 seconds, Fetched: 1 row(s)
hive>
```

. Retrieve all the customers who have account


```
hive> select * from customer , accounts where customer.accid = accounts.accid
    > ;
Query ID = hdoop_20210710111224_b455626a-dcaf-44af-a806-399d54d19fee
Total jobs = 1

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.
2021-07-10 11:12:28    Starting to launch local task to process map join;    maximum memory = 239075328
2021-07-10 11:12:30    Uploaded 1 File to: file:/tmp/hive/java/hdoop/163edaf4-fd7e-4504-aeb1-299db3f44baf/hive_2021-07-10_11-12-24_092_1819826460560943913-1/-local-10004/HashTable-Stage-3/MapJoin-mapfile21--.ha
shtable (326 bytes)
2021-07-10 11:12:30    End of local task; Time Taken: 1.293 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_1625894617298_0005, Tracking URL = http://bharath:8088/proxy/application_1625894617298_0005/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job  -kill job_1625894617298_0005
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0
2021-07-10 11:12:37,494 Stage-3 map = 0%,  reduce = 0%
2021-07-10 11:12:42,637 Stage-3 map = 100%,  reduce = 0%, Cumulative CPU 2.77 sec
MapReduce Total cumulative CPU time: 2 seconds 770 msec
Ended Job = job_1625894617298_0005
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1   Cumulative CPU: 2.77 sec   HDFS Read: 10443 HDFS Write: 246 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 770 msec
OK
1     bharath 34000.0 234     2000-03-01    234   1     22
3     chinmay 12000.0 236     2000-03-04    236   3     20
4     hari    18000.0 237     2000-12-09    237   4     22
Time taken: 21.721 seconds, Fetched: 3 row(s)
hive>
```

Retrieve all the customers whose income is not more than 20000


```
hive> select cname,cust_id ,income from customer where income<=20000;
OK
amith   2       15000.0
chinmay 3       12000.0
hari    4       18000.0
Time taken: 0.208 seconds, Fetched: 3 row(s)
hive>
```

Retrieve all the customers whose dob not between 10-10-1990 and 1-2-1993 group based on cust_id and sort the results in descending order

Retrieve view which contains all the fileds from account and cname in it. Demonstrate update operation it


```
hive> create view details as ( select accid,cust_id,bnkid,cname from accounts,customer where accounts.accid = customer.accid);
FAILED: SemanticException Column accid Found in more than One Tables/Subqueries
hive> create view details as ( select accounts.accid,accounts.cust_id,bnkid,cname from accounts,customer where accounts.accid = customer.accid);
FAILED: SemanticException [Error 10004]: Line 1:64 Invalid table alias or column reference 'bnkid': (possible column names are: accounts.accid, accounts.cust_id, accounts.bnk_id, customer.cust_id, customer.cname
, customer.income, customer.accid, customer.dob)
hive> create view details as ( select accounts.accid,accounts.cust_id,bnk_id,cname from accounts,customer where accounts.accid = customer.accid);
OK
Time taken: 0.862 seconds
hive> select * from details;
Query ID = hdoop_20210710112401_a7ddbc5f-31c3-4d43-bdff-ea6c3ddc112e
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 11:24:07     End of local task; Time Taken: 1.28 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_1625894617298_0008, Tracking URL = http://bharath:8088/proxy/application_1625894617298_0008/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job  -kill job_1625894617298_0008
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0
2021-07-10 11:24:19,158 Stage-3 map = 0%,  reduce = 0%
2021-07-10 11:24:24,312 Stage-3 map = 100%,  reduce = 0%, Cumulative CPU 2.8 sec
MapReduce Total cumulative CPU time: 2 seconds 800 msec
Ended Job = job_1625894617298_0008
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1   Cumulative CPU: 2.8 sec   HDFS Read: 9934 HDFS Write: 171 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 800 msec
OK
234     1       22      bharath
236     3       20      chinmay
237     4       22      hari
Time taken: 25.075 seconds, Fetched: 3 row(s)
hive>
```



```
income              float
accid               int
dob                 date
Time taken: 0.041 seconds, Fetched: 5 row(s)
hive> insert into table customer values(211,'kumar',43000,238,'1992-10-9');
Query ID = hdoop_20210710112830_066a2171-6419-46cc-82d7-aaacf078f5bb
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_1625894617298_0011, Tracking URL = http://bharath:8088/proxy/application_1625894617298_0011/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job  -kill job_1625894617298_0011
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 0
2021-07-10 11:28:37,719 Stage-1 map = 0%,  reduce = 0%
2021-07-10 11:28:42,861 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 2.72 sec
MapReduce Total cumulative CPU time: 2 seconds 720 msec
Ended Job = job_1625894617298_0011
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/organisation.db/customer/.hive-staging_hive_2021-07-10_11-28-30_040_4208558043647418849-1/-ext-10000
Loading data to table organisation.customer
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1   Cumulative CPU: 2.72 sec   HDFS Read: 6097 HDFS Write: 110 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 720 msec
OK
Time taken: 14.202 seconds
hive> select * from details;
Query ID = hdoop_20210710112857_259748c2-4681-4f26-a05f-345d29ac56c3
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 explanation.
2021-07-10 11:29:02     Starting to launch local task to process map join;     maximum memory = 239075328
2021-07-10 11:29:03     Dump the side-table for tag: 0 with group count: 4 into file: file:/tmp/hive/java/hdoop/163edaf4-fd7e-4504-aeb1-299db3f44baf/hive_2021-07-10_11-28-57_743_1325324083298028577-1/-local-1000
4/HashTable-Stage-3/MapJoin-mapfile50-.hashtable
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_1625894617298_0012, Tracking URL = http://bharath:8088/proxy/application_1625894617298_0012/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job  -kill job_1625894617298_0012
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0
2021-07-10 11:29:11,436 Stage-3 map = 0%,  reduce = 0%
2021-07-10 11:29:16,592 Stage-3 map = 100%,  reduce = 0%, Cumulative CPU 2.8 sec
MapReduce Total cumulative CPU time: 2 seconds 800 msec
Ended Job = job_1625894617298_0012
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1   Cumulative CPU: 2.8 sec   HDFS Read: 10066 HDFS Write: 198 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 800 msec
OK
234     1       22      bharath
236     3       20      chinmay
237     4       22      hari
238     8       32      kumar
Time taken: 21.013 seconds, Fetched: 4 row(s)
hive>
```

Loading through csv file

```
hive> show databases;
OK
default
employeedb
organisation
Time taken: 0.651 seconds, Fetched: 3 row(s)
hive> use organisation;
OK
Time taken: 0.031 seconds
hive> show tables;
OK
accounts
bank
customer
details
Time taken: 0.034 seconds, Fetched: 4 row(s)
hive> select * from accounts;
OK
234     1       22
237     4       22
236     3       20
238     8       32
Time taken: 2.367 seconds, Fetched: 4 row(s)
hive> load data local inpath '/home/hdoop/file.csv' into table accounts;
Loading data to table organisation.accounts
OK
Time taken: 1.807 seconds
hive> select * from accounts;
OK
234     1       22
237     4       22
236     3       20
238     8       32
239     9       33
240     10      34
241     11      35
Time taken: 0.183 seconds, Fetched: 7 row(s)
hive> 
```