

**NAME : Chinmai Srivastava**  
**USN : INT18IS046**

Create following tables

Bank(Bank\_id:integer, Bname: string, blocation:string)

Customer(Cust\_id: integer, Cname: string, income:float, accid:int dob:date)

Account(accid int, cust\_id: int,bnk\_id: int);

**Write the HiveQL queries for the above database schema**

**1. Insert 5 records using INSERT command.**

hive> show databases;

OK

default

Time taken: 0.779 seconds, Fetched: 1 row(s)

hive> create database bank\_hive;

OK

Time taken: 0.075 seconds

hive> show databases;

OK

bank\_hive

default

Time taken: 0.064 seconds, Fetched: 2 row(s)

hive> use bank\_hive;

OK

Time taken: 0.071 seconds

hive> create table bank(bank\_id integer,bname string,blocation string) row format delimited fields terminated by ',';

OK

Time taken: 0.946 seconds

hive>

> insert into table bank

values(001,'sbi','bangalore'),(002,'hdfc','mumbai'),(003,'axis','punjab'),(004,'syndicate','mysore'),(005,'axis','chennai');

Query ID = hdoop\_20210705215945\_91b2e084-743a-461a-85cf-672f60de7d69

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_1625493478645_0001, Tracking URL = http://hadoop-
VirtualBox:8088/proxy/application_1625493478645_0001/
Kill Command = /home/hadoop/hadoop-3.2.1/bin/mapred job -kill
job_1625493478645_0001
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2021-07-05 22:00:08,130 Stage-1 map = 0%, reduce = 0%
2021-07-05 22:00:16,883 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.23 sec
2021-07-05 22:00:24,381 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 4.0 sec
MapReduce Total cumulative CPU time: 4 seconds 0 msec
Ended Job = job_1625493478645_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/bank_hive.db/bank/.hive-
staging_hive_2021-07-05_21-59-45_097_5547249773166002765-1/-ext-10000
Loading data to table bank_hive.bank
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.0 sec HDFS Read: 17204 HDFS
Write: 406 SUCCESS
Total MapReduce CPU Time Spent: 4 seconds 0 msec
OK
Time taken: 43.036 seconds
hive> select * from bank;
OK
1      sbi      bangalore
2      hdfc     mumbai
3      axis     punjab
4      syndicate mysore
5      axis     chennai
Time taken: 0.34 seconds, Fetched: 5 row(s)
hive>
> create table customer(cust_id integer,cname string,income float,acc_id int,dob
date) row format delimited fields terminated by ',';
OK
Time taken: 0.155 seconds
hive> insert into table customer values(101,'Bianca',40000,201,'1991-11-
10'),(102,'Hani',15000,202,'1995-05-12'),(103,'Lavanya',18000,203,'1999-03-
08'),(104,'Anita',25000,204,'1992-07-09'),(105,'Tanya',16000,205,'2000-11-17');
Query ID = hadoop_20210705221018_2b376a5b-6c45-4667-b44f-84ad0e8aeb16
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\_1625493478645\_0002, Tracking URL = http://hadoop-VirtualBox:8088/proxy/application\_1625493478645\_0002/  
 Kill Command = /home/hadoop/hadoop-3.2.1/bin/mapred job -kill job\_1625493478645\_0002  
 Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 0  
 2021-07-05 22:10:30,187 Stage-1 map = 0%, reduce = 0%  
 2021-07-05 22:10:38,679 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.99 sec  
 MapReduce Total cumulative CPU time: 1 seconds 990 msec  
 Ended Job = job\_1625493478645\_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/bank\_hive.db/customer/.hive-staging\_hive\_2021-07-05\_22-10-18\_758\_660404485710494493-1/-ext-10000  
 Loading data to table bank\_hive.customer  
 MapReduce Jobs Launched:  
 Stage-Stage-1: Map: 1 Cumulative CPU: 1.99 sec HDFS Read: 6556 HDFS Write: 242  
 SUCCESS  
 Total MapReduce CPU Time Spent: 1 seconds 990 msec  
 OK  
 Time taken: 21.459 seconds  
 hive> select \* from customer;  
 OK  

101	Bianca	40000.0	201	1991-11-10
102	Hani	15000.0	202	1995-05-12
103	Lavanya	18000.0	203	1999-03-08
104	Anita	25000.0	204	1992-07-09
105	Tanya	16000.0	205	2000-11-17

 Time taken: 0.233 seconds, Fetched: 5 row(s)  
 hive>  
 > create table account(acc\_id integer,cust\_id integer,bank\_id integer) row format delimited fields terminated by ',';  
 OK  
 Time taken: 0.116 seconds  
 hive> insert into table account  
 values(201,101,003),(202,102,001),(203,103,004),(204,104,005),(205,105,002);  
 Query ID = hdoop\_20210705221427\_feb99943-e0b3-4912-831b-9c3476c7dedc  
 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_1625493478645_0003, Tracking URL = http://hadoop-
VirtualBox:8088/proxy/application_1625493478645_0003/
Kill Command = /home/hadoop/hadoop-3.2.1/bin/mapred job -kill
job_1625493478645_0003
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2021-07-05 22:14:38,236 Stage-1 map = 0%, reduce = 0%
2021-07-05 22:14:46,698 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.1 sec
2021-07-05 22:14:55,120 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 3.85 sec
MapReduce Total cumulative CPU time: 3 seconds 850 msec
Ended Job = job_1625493478645_0003
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/bank_hive.db/account/.hive-
staging_hive_2021-07-05_22-14-27_096_2649192857733927476-1/-ext-10000
Loading data to table bank_hive.account
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 3.85 sec HDFS Read: 15906 HDFS
Write: 389 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 850 msec
OK
Time taken: 29.675 seconds
hive> select * from account;
OK
201    101    3
202    102    1
203    103    4
204    104    5
205    105    2
Time taken: 0.269 seconds, Fetched: 5 row(s)
hive>
> show tables;
OK
account
bank
customer
Time taken: 0.046 seconds, Fetched: 3 row(s)

```

## 1. Rename the table Account to Accounts.

```
hive>  
  > alter table account rename to accounts;  
OK  
Time taken: 0.374 seconds  
hive> show tables;  
OK  
accounts  
bank  
customer  
Time taken: 0.038 seconds, Fetched: 3 row(s)
```

```
Time taken: 29.675 seconds  
hive> select * from account;  
OK  
201      101      3  
202      102      1  
203      103      4  
204      104      5  
205      105      2  
Time taken: 0.269 seconds, Fetched: 5 row(s)  
hive>  
  > show tables;  
OK  
account  
bank  
customer  
Time taken: 0.046 seconds, Fetched: 3 row(s)  
hive>  
  > alter table account rename to accounts;  
OK  
Time taken: 0.374 seconds  
hive> show tables;  
OK  
accounts  
bank  
customer  
Time taken: 0.038 seconds, Fetched: 3 row(s)  
hive>
```

## 2. Rename the column name "blocation" to "location".

```
hive>  
  > desc bank;  
OK  
bank_id      int  
bname        string  
blocation    string  
Time taken: 0.099 seconds, Fetched: 3 row(s)  
hive>  
  > alter table bank change blocation location string;  
OK  
Time taken: 0.271 seconds  
hive> desc bank;
```

OK

bank_id	int
bname	string
location	string

Time taken: 0.108 seconds, Fetched: 3 row(s)

```
hive>
> desc bank;
OK
bank_id          int
bname            string
blocation        string
Time taken: 0.099 seconds, Fetched: 3 row(s)
hive>
> alter table bank change blocation location string;
OK
Time taken: 0.271 seconds
hive> desc bank;
OK
bank_id          int
bname            string
location         string
Time taken: 0.108 seconds, Fetched: 3 row(s)
hive> █
```

### 3. Retrieve all the customers who have account in axis bank

```
hive> select * from customer c,bank b,accounts a where c.cust_id=a.cust_id and
b.bank_id=a.bank_id and b.bname='axis';
```

No Stats for bank\_hive@customer, Columns: income, dob, cname, cust\_id, acc\_id

No Stats for bank\_hive@bank, Columns: location

Query ID = hdoop\_20210705232651\_e4e33a79-9a9b-4f9e-b9cb-c097a3d978cd

Total jobs = 1

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\_1625493478645\_0004, Tracking URL = [http://hadoop-VirtualBox:8088/proxy/application\\_1625493478645\\_0004/](http://hadoop-VirtualBox:8088/proxy/application_1625493478645_0004/)

Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job -kill
job\_1625493478645\_0004

Hadoop job information for Stage-5: number of mappers: 1; number of reducers: 0

2021-07-05 23:27:18,188 Stage-5 map = 0%, reduce = 0%

2021-07-05 23:27:26,692 Stage-5 map = 100%, reduce = 0%, Cumulative CPU 2.37 sec

MapReduce Total cumulative CPU time: 2 seconds 370 msec

Ended Job = job\_1625493478645\_0004

MapReduce Jobs Launched:

Stage-Stage-5: Map: 1 Cumulative CPU: 2.37 sec HDFS Read: 13481 HDFS Write: 227
SUCCESS

Total MapReduce CPU Time Spent: 2 seconds 370 msec

OK

101	Bianca	40000.0	201	1991-11-10	3	axis	punjab	201	101
	3								
104	Anita	25000.0	204	1992-07-09	5	axis	chennai	204	104
	5								

Time taken: 36.922 seconds, Fetched: 2 row(s)

```
hive> select * from customer c,bank b,accounts a where c.cust_id=a.cust_id and
b.bank_id=a.bank_id and b.bname='axis';
No Stats for bank_hive@customer, Columns: income, dob, cname, cust_id, acc_id
No Stats for bank_hive@bank, Columns: location
Query ID = hdoop_20210705232651_e4e33a79-9a9b-4f9e-b9cb-c097a3d978cd
Total jobs = 1
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_1625493478645_0004, Tracking URL = http://hadoop-VirtualBox:
8088/proxy/application_1625493478645_0004/
Kill Command = /home/hadoop/hadoop-3.2.1/bin/mapred job -kill job_1625493478645
_0004
Hadoop job information for Stage-5: number of mappers: 1; number of reducers: 0
2021-07-05 23:27:18,188 Stage-5 map = 0%, reduce = 0%
2021-07-05 23:27:26,692 Stage-5 map = 100%, reduce = 0%, Cumulative CPU 2.37 s
ec
MapReduce Total cumulative CPU time: 2 seconds 370 msec
Ended Job = job_1625493478645_0004
```

```
No Stats for bank_hive@customer, Columns: income, dob, cname, cust_id, acc_id
No Stats for bank_hive@bank, Columns: location
Query ID = hdoop_20210705232651_e4e33a79-9a9b-4f9e-b9cb-c097a3d978cd
Total jobs = 1
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_1625493478645_0004, Tracking URL = http://hadoop-VirtualBox:
8088/proxy/application_1625493478645_0004/
Kill Command = /home/hadoop/hadoop-3.2.1/bin/mapred job -kill job_1625493478645
_0004
Hadoop job information for Stage-5: number of mappers: 1; number of reducers: 0
2021-07-05 23:27:18,188 Stage-5 map = 0%, reduce = 0%
2021-07-05 23:27:26,692 Stage-5 map = 100%, reduce = 0%, Cumulative CPU 2.37 s
ec
MapReduce Total cumulative CPU time: 2 seconds 370 msec
Ended Job = job_1625493478645_0004
MapReduce Jobs Launched:
Stage-Stage-5: Map: 1 Cumulative CPU: 2.37 sec HDFS Read: 13481 HDFS Write:
227 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 370 msec
OK
101 Bianca 40000.0 201 1991-11-10 3 axis punjab 201 1
01 3
104 Anita 25000.0 204 1992-07-09 5 axis chennai 204 1
04 5
Time taken: 36.922 seconds, Fetched: 2 row(s)
hive>
```

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

hive>

> select max(income) as max\_income,min(income) as min\_income from customer;

Query ID = hdoop\_20210705233038\_aa2a98e8-f0c6-4da3-9364-63d004e8bac3

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\_1625493478645\_0005, Tracking URL = http://hadoop-

VirtualBox:8088/proxy/application\_1625493478645\_0005/

Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job -kill

job\_1625493478645\_0005

Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1

2021-07-05 23:30:49,780 Stage-1 map = 0%, reduce = 0%

2021-07-05 23:30:57,192 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.58 sec

2021-07-05 23:31:05,568 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 3.47 sec

MapReduce Total cumulative CPU time: 3 seconds 470 msec

Ended Job = job\_1625493478645\_0005

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 3.47 sec HDFS Read: 14368 HDFS

Write: 115 SUCCESS

Total MapReduce CPU Time Spent: 3 seconds 470 msec

OK

40000.0 15000.0

Time taken: 27.895 seconds, Fetched: 1 row(s)

```
hive>
> select max(income) as max_income,min(income) as min_income from customer;

Query ID = hdoop_20210705233038_aa2a98e8-f0c6-4da3-9364-63d004e8bac3
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_1625493478645_0005, Tracking URL = http://hadoop-VirtualBox:
8088/proxy/application_1625493478645_0005/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job -kill job_1625493478645
_0005
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2021-07-05 23:30:49,780 Stage-1 map = 0%, reduce = 0%
2021-07-05 23:30:57,192 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.58 s
ec
2021-07-05 23:31:05,568 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 3.47
sec
MapReduce Total cumulative CPU time: 3 seconds 470 msec
Ended Job = job_1625493478645_0005
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 3.47 sec HDFS Read: 14368
```



```

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_1625493478645_0005, Tracking URL = http://hadoop-VirtualBox:
8088/proxy/application_1625493478645_0005/
Kill Command = /home/hadoop/hadoop-3.2.1/bin/mapred job -kill job_1625493478645
_0005
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2021-07-05 23:30:49,780 Stage-1 map = 0%, reduce = 0%
2021-07-05 23:30:57,192 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.58 s
ec
2021-07-05 23:31:05,568 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 3.47
sec
MapReduce Total cumulative CPU time: 3 seconds 470 msec
Ended Job = job_1625493478645_0005
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 3.47 sec HDFS Read: 14368
HDFS Write: 115 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 470 msec
OK
40000.0 15000.0
Time taken: 27.895 seconds, Fetched: 1 row(s)
hive>

```

## 5. Retrieve all the customers who have account

hive>

> select \* from customer where acc\_id IS NOT NULL;

OK

101	Bianca	40000.0	201	1991-11-10
102	Hani	15000.0	202	1995-05-12
103	Lavanya	18000.0	203	1999-03-08
104	Anita	25000.0	204	1992-07-09
105	Tanya	16000.0	205	2000-11-17
101	Bianca	40000.0	201	1991-11-10
102	Hani	15000.0	202	1995-05-12
103	Lavanya	18000.0	203	1999-03-08
104	Anita	25000.0	204	1992-07-09
105	Tanya	16000.0	205	2000-11-17
106	Sam	30000.0	206	1995-10-17

Time taken: 0.252 seconds, Fetched: 11 row(s)

```
hive>
> select * from customer where acc_id IS NOT NULL;
OK
101    Bianca  40000.0 201      1991-11-10
102    Hani    15000.0 202      1995-05-12
103    Lavanya 18000.0 203      1999-03-08
104    Anita   25000.0 204      1992-07-09
105    Tanya   16000.0 205      2000-11-17
101    Bianca  40000.0 201      1991-11-10
102    Hani    15000.0 202      1995-05-12
103    Lavanya 18000.0 203      1999-03-08
104    Anita   25000.0 204      1992-07-09
105    Tanya   16000.0 205      2000-11-17
106    Sam     30000.0 206      1995-10-17
Time taken: 0.252 seconds, Fetched: 11 row(s)
```

## 6. Retrieve all the customers whose income is not more than 20000

```
hive>
> select * from customer where income<=20000;
OK
102    Hani  15000.0      202    1995-05-12
103    Lavanya 18000.0      203    1999-03-08
105    Tanya 16000.0      205    2000-11-17
Time taken: 0.268 seconds, Fetched: 3 row(s)
```

```
hive>
> select * from customer where income<=20000;
OK
102    Hani    15000.0 202      1995-05-12
103    Lavanya 18000.0 203      1999-03-08
105    Tanya   16000.0 205      2000-11-17
Time taken: 0.268 seconds, Fetched: 3 row(s)
hive>
```

## 7. 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

```
hive>
>
> select * from customer where dob not in(select dob from customer where dob
between '1990-10-10' and '1993-02-01') order by cust_id desc;
Warning: Map Join MAPJOIN[46][bigTable=?] in task 'Stage-8:MAPRED' is a cross
product
Query ID = hdoop_20210705234450_abff544a-0cd5-4eda-85d3-c0262b1ee5a9
Total jobs = 6
Launching Job 1 out of 6
```

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\_1625493478645\_0006, Tracking URL = http://hadoop-VirtualBox:8088/proxy/application\_1625493478645\_0006/

Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job -kill job\_1625493478645\_0006

Hadoop job information for Stage-4: number of mappers: 1; number of reducers: 1

2021-07-05 23:45:01,515 Stage-4 map = 0%, reduce = 0%

2021-07-05 23:45:09,238 Stage-4 map = 100%, reduce = 0%, Cumulative CPU 2.3 sec

2021-07-05 23:45:17,590 Stage-4 map = 100%, reduce = 100%, Cumulative CPU 4.07 sec

MapReduce Total cumulative CPU time: 4 seconds 70 msec

Ended Job = job\_1625493478645\_0006

Launching Job 2 out of 6

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\_1625493478645\_0007, Tracking URL = http://hadoop-VirtualBox:8088/proxy/application\_1625493478645\_0007/

Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job -kill job\_1625493478645\_0007

Hadoop job information for Stage-5: number of mappers: 1; number of reducers: 1

2021-07-05 23:45:34,301 Stage-5 map = 0%, reduce = 0%

2021-07-05 23:45:41,733 Stage-5 map = 100%, reduce = 0%, Cumulative CPU 2.14 sec

2021-07-05 23:45:50,089 Stage-5 map = 100%, reduce = 100%, Cumulative CPU 3.94 sec

MapReduce Total cumulative CPU time: 3 seconds 940 msec

Ended Job = job\_1625493478645\_0007

Execution completed successfully

MapredLocal task succeeded

Launching Job 3 out of 6

Number of reduce tasks is set to 0 since there's no reduce operator

Starting Job = job\_1625493478645\_0008, Tracking URL = http://hadoop-VirtualBox:8088/proxy/application\_1625493478645\_0008/

Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job -kill job\_1625493478645\_0008

Hadoop job information for Stage-8: number of mappers: 1; number of reducers: 0  
2021-07-05 23:46:12,231 Stage-8 map = 0%, reduce = 0%  
2021-07-05 23:46:19,593 Stage-8 map = 100%, reduce = 0%, Cumulative CPU 1.68 sec  
MapReduce Total cumulative CPU time: 1 seconds 680 msec  
Ended Job = job\_1625493478645\_0008  
Stage-9 is selected by condition resolver.  
Stage-2 is filtered out by condition resolver.  
2021-07-05 23:46:32 Dump the side-table for tag: 1 with group count: 2 into file:  
file:/tmp/hive/java/hadoop/c2662ef9-49c1-43ab-8d18-059e0fc52024/hive\_2021-07-05\_23-44-50\_492\_5772229897397811597-1/-local-10008/HashTable-Stage-6/MapJoin-mapfile21--.hashtable  
Execution completed successfully  
MapredLocal task succeeded  
Launching Job 5 out of 6  
Number of reduce tasks is set to 0 since there's no reduce operator  
Starting Job = job\_1625493478645\_0009, Tracking URL = http://hadoop-VirtualBox:8088/proxy/application\_1625493478645\_0009/  
Kill Command = /home/hadoop/hadoop-3.2.1/bin/mapred job -kill job\_1625493478645\_0009  
Hadoop job information for Stage-6: number of mappers: 1; number of reducers: 0  
2021-07-05 23:46:43,332 Stage-6 map = 0%, reduce = 0%  
2021-07-05 23:46:51,738 Stage-6 map = 100%, reduce = 0%, Cumulative CPU 2.55 sec  
MapReduce Total cumulative CPU time: 2 seconds 550 msec  
Ended Job = job\_1625493478645\_0009  
Launching Job 6 out of 6  
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\_1625493478645\_0010, Tracking URL = http://hadoop-VirtualBox:8088/proxy/application\_1625493478645\_0010/  
Kill Command = /home/hadoop/hadoop-3.2.1/bin/mapred job -kill job\_1625493478645\_0010  
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 1  
2021-07-05 23:47:08,397 Stage-3 map = 0%, reduce = 0%  
2021-07-05 23:47:15,716 Stage-3 map = 100%, reduce = 0%, Cumulative CPU 1.62 sec  
2021-07-05 23:47:24,043 Stage-3 map = 100%, reduce = 100%, Cumulative CPU 3.35 sec  
MapReduce Total cumulative CPU time: 3 seconds 350 msec  
Ended Job = job\_1625493478645\_0010  
MapReduce Jobs Launched:

Stage-Stage-4: Map: 1 Reduce: 1 Cumulative CPU: 4.07 sec HDFS Read: 14468 HDFS Write: 115 SUCCESS

Stage-Stage-5: Map: 1 Reduce: 1 Cumulative CPU: 3.94 sec HDFS Read: 13072 HDFS Write: 138 SUCCESS

Stage-Stage-8: Map: 1 Cumulative CPU: 1.68 sec HDFS Read: 6884 HDFS Write: 273 SUCCESS

Stage-Stage-6: Map: 1 Cumulative CPU: 2.55 sec HDFS Read: 9094 HDFS Write: 196 SUCCESS

Stage-Stage-3: Map: 1 Reduce: 1 Cumulative CPU: 3.35 sec HDFS Read: 8932 HDFS Write: 223 SUCCESS

Total MapReduce CPU Time Spent: 15 seconds 590 msec

OK

105	Tanya	16000.0	205	2000-11-17
103	Lavanya	18000.0	203	1999-03-08
102	Hani	15000.0	202	1995-05-12

Time taken: 154.685 seconds, Fetched: 3 row(s)

```
hive>
>
> select * from customer where dob not in(select dob from customer where do
b between '1990-10-10' and '1993-02-01') order by cust_id desc;
Warning: Map Join MAPJOIN[46][bigTable=?] in task 'Stage-8:MAPRED' is a cross p
roduct
Query ID = hdoop_20210705234450_abff544a-0cd5-4eda-85d3-c0262b1ee5a9
Total jobs = 6
Launching Job 1 out of 6
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_1625493478645_0006, Tracking URL = http://hadoop-VirtualBox:8088/proxy/application\_1625493478645\_0006/
Kill Command = /home/hadoop/hadoop-3.2.1/bin/mapred job -kill job_1625493478645_0006
Hadoop job information for Stage-4: number of mappers: 1; number of reducers: 1
2021-07-05 23:45:01,515 Stage-4 map = 0%, reduce = 0%
```



```

8088/proxy/application_1625493478645_0010/
Kill Command = /home/hadoop/hadoop-3.2.1/bin/mapred job -kill job_1625493478645_0010
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 1
2021-07-05 23:47:08,397 Stage-3 map = 0%, reduce = 0%
2021-07-05 23:47:15,716 Stage-3 map = 100%, reduce = 0%, Cumulative CPU 1.62 sec
2021-07-05 23:47:24,043 Stage-3 map = 100%, reduce = 100%, Cumulative CPU 3.35 sec
MapReduce Total cumulative CPU time: 3 seconds 350 msec
Ended Job = job_1625493478645_0010
MapReduce Jobs Launched:
Stage-Stage-4: Map: 1 Reduce: 1 Cumulative CPU: 4.07 sec HDFS Read: 14468
HDFS Write: 115 SUCCESS
Stage-Stage-5: Map: 1 Reduce: 1 Cumulative CPU: 3.94 sec HDFS Read: 13072
HDFS Write: 138 SUCCESS
Stage-Stage-8: Map: 1 Cumulative CPU: 1.68 sec HDFS Read: 6884 HDFS Write:
273 SUCCESS
Stage-Stage-6: Map: 1 Cumulative CPU: 2.55 sec HDFS Read: 9094 HDFS Write:
196 SUCCESS
Stage-Stage-3: Map: 1 Reduce: 1 Cumulative CPU: 3.35 sec HDFS Read: 8932 H
DFS Write: 223 SUCCESS
Total MapReduce CPU Time Spent: 15 seconds 590 msec
OK
105 Tanya 16000.0 205 2000-11-17
103 Lavanya 18000.0 203 1999-03-08
102 Hani 15000.0 202 1995-05-12
Time taken: 154.685 seconds, Fetched: 3 row(s)
hive>

```

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

hive>

>

> create view cust\_acc as select c.cname,a.acc\_id,a.cust\_id,a.bank\_id from customer c,accounts a where c.cust\_id=a.cust\_id;

OK

Time taken: 0.389 seconds

hive> select \* from cust\_acc;

Query ID = hdoop\_20210705235252\_feef2ca6-ac7f-476d-878a-7c682212de47

Total jobs = 1

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\_1625493478645\_0011, Tracking URL = http://hadoop-VirtualBox:8088/proxy/application\_1625493478645\_0011/

Kill Command = /home/hadoop/hadoop-3.2.1/bin/mapred job -kill job\_1625493478645\_0011

Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0

2021-07-05 23:53:15,711 Stage-3 map = 0%, reduce = 0%

2021-07-05 23:53:24,109 Stage-3 map = 100%, reduce = 0%, Cumulative CPU 2.25 sec

MapReduce Total cumulative CPU time: 2 seconds 250 msec

Ended Job = job\_1625493478645\_0011

MapReduce Jobs Launched:

Stage-Stage-3: Map: 1 Cumulative CPU: 2.25 sec HDFS Read: 10208 HDFS Write: 229  
SUCCESS

Total MapReduce CPU Time Spent: 2 seconds 250 msec

OK

Bianca201 101 3

Hani 202 102 1

Lavanya 203 103 4

Anita 204 104 5

Tanya 205 105 2

Time taken: 34.004 seconds, Fetched: 5 row(s)

hive>

> update cust\_acc set cname="Sam" where bank\_id=5;

FAILED: SemanticException [Error 10294]: Attempt to do update or delete using transaction manager that does not support these operations.

hive>

hive> select \* from cust\_acc;

Query ID = hdoop\_20210707171613\_d105a99c-7a38-425d-999d-4ff3e5e54fa1

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]

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\_1625642025578\_0004, Tracking URL = http://hadoop-VirtualBox:8088/proxy/application\_1625642025578\_0004/

Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job -kill job\_1625642025578\_0004

Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0

2021-07-07 17:16:57,466 Stage-3 map = 0%, reduce = 0%

2021-07-07 17:17:13,717 Stage-3 map = 100%, reduce = 0%, Cumulative CPU 3.15 sec

MapReduce Total cumulative CPU time: 3 seconds 150 msec

Ended Job = job\_1625642025578\_0004

MapReduce Jobs Launched:

Stage-Stage-3: Map: 1 Cumulative CPU: 3.15 sec HDFS Read: 10349 HDFS Write: 371  
SUCCESS

Total MapReduce CPU Time Spent: 3 seconds 150 msec

OK

Bianca	201	101	3	
Hani	202	102	1	
Lavanya	203	103	4	
Anita	204	104	5	
Tanya	205	105	2	
Bianca	201	101	3	
Hani	202	102	1	
Lavanya	203	103	4	
Anita	204	104	5	
Tanya	205	105	2	

Time taken: 62.049 seconds, Fetched: 10 row(s)

hive>

> insert into table customer values(106,'Sam',30000,206,'1995-10-17');

Query ID = hdoop\_20210707172047\_c34439eb-bd42-485b-bc8f-f4e88f833663

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\_1625642025578\_0005, Tracking URL = http://hadoop-

VirtualBox:8088/proxy/application\_1625642025578\_0005/

Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job -kill

job\_1625642025578\_0005

Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 0

2021-07-07 17:20:59,949 Stage-1 map = 0%, reduce = 0%

2021-07-07 17:21:08,427 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.15 sec

MapReduce Total cumulative CPU time: 2 seconds 150 msec

Ended Job = job\_1625642025578\_0005

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/bank\_hive.db/customer/.hive-

staging\_hive\_2021-07-07\_17-20-47\_862\_2750762556130156905-1/-ext-10000

Loading data to table bank\_hive.customer

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Cumulative CPU: 2.15 sec HDFS Read: 6124 HDFS Write: 105

SUCCESS

Total MapReduce CPU Time Spent: 2 seconds 150 msec

OK

Time taken: 23.087 seconds

hive> insert into table account values(206,103,002);

Query ID = hdoop\_20210707172136\_b391e9ff-0791-4d67-8d74-8138a53b6f61

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\_1625642025578\_0006, Tracking URL = http://hadoop-VirtualBox:8088/proxy/application\_1625642025578\_0006/  
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job -kill job\_1625642025578\_0006  
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1  
2021-07-07 17:21:48,930 Stage-1 map = 0%, reduce = 0%  
2021-07-07 17:21:57,485 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.22 sec  
2021-07-07 17:22:05,852 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 4.04 sec  
MapReduce Total cumulative CPU time: 4 seconds 40 msec  
Ended Job = job\_1625642025578\_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/bank\_hive.db/account/.hive-staging\_hive\_2021-07-07\_17-21-36\_876\_6245912131836166744-1/-ext-10000  
Loading data to table bank\_hive.account  
MapReduce Jobs Launched:  
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.04 sec HDFS Read: 15619 HDFS Write: 280 SUCCESS  
Total MapReduce CPU Time Spent: 4 seconds 40 msec  
OK  
Time taken: 31.664 seconds  
hive> select \* from cust\_acc;  
Query ID = hdoop\_20210707172213\_f1281b42-d2c1-42e2-b733-25f9b83bdb60  
Total jobs = 1  
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\_1625642025578\_0007, Tracking URL = http://hadoop-VirtualBox:8088/proxy/application\_1625642025578\_0007/  
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job -kill job\_1625642025578\_0007  
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0  
2021-07-07 17:22:37,423 Stage-3 map = 0%, reduce = 0%

2021-07-07 17:22:45,773 Stage-3 map = 100%, reduce = 0%, Cumulative CPU 2.28 sec

MapReduce Total cumulative CPU time: 2 seconds 280 msec

Ended Job = job\_1625642025578\_0007

MapReduce Jobs Launched:

Stage-Stage-3: Map: 1 Cumulative CPU: 2.28 sec HDFS Read: 10487 HDFS Write: 431  
SUCCESS

Total MapReduce CPU Time Spent: 2 seconds 280 msec

OK

Bianca	201	101	3
Hani	202	102	1
Lavanya	203	103	4
Lavanya	206	103	2
Anita	204	104	5
Tanya	205	105	2
Bianca	201	101	3
Hani	202	102	1
Lavanya	203	103	4
Lavanya	206	103	2
Anita	204	104	5
Tanya	205	105	2

Time taken: 33.341 seconds, Fetched: 12 row(s)

hive>

```
hive>
> create view cust_acc as select c.cname,a.acc_id,a.cust_id,a.bank_id from
customer c,account a where c.cust_id=a.cust_id;
OK
Time taken: 0.659 seconds
hive> select * from cust_acc;
Query ID = hdoop_20210707171613_d105a99c-7a38-425d-999d-4ff3e5e54fa1
Total jobs = 1
SLF4J: Found binding in [jar:file:/home/hdoop/apache-hive-3.1.2-bin/lib/log4j-s
lf4j-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]
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_1625642025578_0004, Tracking URL = http://hadoop-VirtualBox:
8088/proxy/application_1625642025578_0004/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job -kill job_1625642025578
_0004
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0
2021-07-07 17:16:57,466 Stage-3 map = 0%, reduce = 0%
2021-07-07 17:17:13,717 Stage-3 map = 100%, reduce = 0%, Cumulative CPU 3.15 s
ec
MapReduce Total cumulative CPU time: 3 seconds 150 msec
Ended Job = job_1625642025578_0004
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1 Cumulative CPU: 3.15 sec HDFS Read: 10349 HDFS Write:
```

```

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_1625642025578_0004, Tracking URL = http://hadoop-VirtualBox:
8088/proxy/application_1625642025578_0004/
Kill Command = /home/hadoop/hadoop-3.2.1/bin/mapred job -kill job_1625642025578
_0004
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0
2021-07-07 17:16:57,466 Stage-3 map = 0%, reduce = 0%
2021-07-07 17:17:13,717 Stage-3 map = 100%, reduce = 0%, Cumulative CPU 3.15 s
ec
MapReduce Total cumulative CPU time: 3 seconds 150 msec
Ended Job = job_1625642025578_0004
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1 Cumulative CPU: 3.15 sec HDFS Read: 10349 HDFS Write:
371 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 150 msec
OK
Bianca 201 101 3
Hani 202 102 1
Lavanya 203 103 4
Anita 204 104 5
Tanya 205 105 2
Bianca 201 101 3
Hani 202 102 1
Lavanya 203 103 4
Anita 204 104 5
Tanya 205 105 2
Time taken: 62.049 seconds, Fetched: 10 row(s)

```

```

hive>
> insert into table customer values(106,'Sam',30000,206,'1995-10-17');
Query ID = hdoop_20210707172047_c34439eb-bd42-485b-bc8f-f4e88f833663
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_1625642025578_0005, Tracking URL = http://hadoop-VirtualBox:
8088/proxy/application_1625642025578_0005/
Kill Command = /home/hadoop/hadoop-3.2.1/bin/mapred job -kill job_1625642025578
_0005
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 0
2021-07-07 17:20:59,949 Stage-1 map = 0%, reduce = 0%
2021-07-07 17:21:08,427 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.15 s
ec
MapReduce Total cumulative CPU time: 2 seconds 150 msec
Ended Job = job_1625642025578_0005
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/bank_hive.db
/customer/.hive-staging_hive_2021-07-07_17-20-47_862_2750762556130156905-1/-ext
-10000
Loading data to table bank_hive.customer
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Cumulative CPU: 2.15 sec HDFS Read: 6124 HDFS Write:
105 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 150 msec
OK
Time taken: 23.087 seconds

```



```

hive> insert into table account values(206,103,002);
Query ID = hdoop_20210707172136_b391e9ff-0791-4d67-8d74-8138a53b6f61
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_1625642025578_0006, Tracking URL = http://hadoop-VirtualBox:
8088/proxy/application_1625642025578_0006/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job -kill job_1625642025578
_0006
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2021-07-07 17:21:48,930 Stage-1 map = 0%, reduce = 0%
2021-07-07 17:21:57,485 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.22 s
ec
2021-07-07 17:22:05,852 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 4.04
sec
MapReduce Total cumulative CPU time: 4 seconds 40 msec
Ended Job = job_1625642025578_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/bank_hive.db
/account/.hive-staging_hive_2021-07-07_17-21-36_876_6245912131836166744-1/-ext-
10000

```

Bianca	201	101	3
Hani	202	102	1
Lavanya	203	103	4
Lavanya	206	103	2
Anita	204	104	5
Tanya	205	105	2
Bianca	201	101	3
Hani	202	102	1
Lavanya	203	103	4
Lavanya	206	103	2
Anita	204	104	5
Tanya	205	105	2

Time taken: 33.341 seconds, Fetched: 12 row(s)

```
hive>
```