

## Experiment 4

(1)

BDI Experiment 4

DATE:

Kareena Shah

60004210243

C'22

Aim : Execute HIVE commands to load , insert , retrieve ,  
update or delete data in the tables

Theory :

Hive is a data warehouse infrastructure tool to process  
structured data in Hadoop

It resides on top of Hadoop to summarize Big Data & makes  
querying & analyzing easy

Features :

- (1) It stores schema in a database & processed data into HDFS
- (2) It is designed for OLAP
- (3) It provides SQL type language for querying called HiveQL or  
HQL
- (4) It is familiar , fast , scalable & extensible

Commands :

- show databases
- create database database name
- use database name

Create a table (optional clauses)

IF NOT EXISTS

EXTERNAL

TEMPORARY

DATE:

ROW FORMAT

FIELDS TERMINATED BY

PARTITION BY

CLUSTERED BY

LOCATION

- Insert into table\_name ... values

- Order By clause

- Group By clause

- Rename table

alter table table\_name rename to new table\_name

- Add new column in table

alter table stud add columns (column\_name column datatype)

- Aggregate functions

SUM

MAX

MIN

AVG

COUNT

- Arithmetic Operators & Logical Operators

AND (&&)

+

/

^

OR (||)

-

%

~

NOT (!)

\*

|

&

- Joins

Left Outer Join

Inner Join

Right Outer Join

Full Outer Join

DATE:

### - Views

Create view if not exists table\_name\_view

as select \* from table\_name where condition = "value";

drop view if exists table\_name view;

Conclusion : Thus, we implemented HIVE commands successfully.

FOR EDUCATIONAL USE

```
hive> show databases;  
OK  
default  
foodmart  
xademo  
Time taken: 1.771 seconds, Fetched: 3 row(s)
```

```
hive> create database Daksh;
OK
Time taken: 6.58 seconds
```

```
hive> use Daksh;
OK
Time taken: 0.316 seconds
```

```
hive> create table employee(id int, name varchar(20), dept varchar(20), salary int)
> PARTITIONED BY (load_date date)
> CLUSTERED BY (id) INTO 3 BUCKETS
> STORED AS ORC TBLPROPERTIES('transactional'='true');
OK
Time taken: 1.336 seconds
```

```
hive> insert into employee partition (load_date='2024-02-14') values (001,'Seon','Business Analyst',50000);
Query ID = root_20240207044803_82e9ca32-dc3e-468f-afca-79be37235491
Total jobs = 1
Launching Job 1 out of 1
Tez session was closed. Reopening...
Session re-established.
```

Status: Running (Executing on YARN cluster with App id application\_1707276530290\_0002)

| VERTICES    | STATUS    | TOTAL | COMPLETED | RUNNING | PENDING | FAILED | KILLED |
|-------------|-----------|-------|-----------|---------|---------|--------|--------|
| Map 1 ..... | SUCCEEDED | 1     | 1         | 0       | 0       | 0      | 0      |

VERTICES: 01/01 [=====>>] 100% ELAPSED TIME: 15.94 s

```
Loading data to table daksh.employee partition (load_date=2024-02-14)
Partition daksh.employee{load_date=2024-02-14} stats: [numFiles=1, numRows=0, totalSize=860, rawDataSize=0]
OK
Time taken: 38.563 seconds
```

```
hive> insert into employee partition (load_date='2024-02-13') values (002,'Rohan','Trainee',25000);
Query ID = root_20240207045106_1251c0c6-0bc1-4ada-8ad2-77775d5893a4
Total jobs = 1
Launching Job 1 out of 1
```

Status: Running (Executing on YARN cluster with App id application\_1707276530290\_0002)

| VERTICES    | STATUS    | TOTAL | COMPLETED | RUNNING | PENDING | FAILED | KILLED |
|-------------|-----------|-------|-----------|---------|---------|--------|--------|
| Map 1 ..... | SUCCEEDED | 1     | 1         | 0       | 0       | 0      | 0      |

VERTICES: 01/01 [=====>>] 100% ELAPSED TIME: 8.85 s

```
Loading data to table daksh.employee partition (load_date=2024-02-13)
Partition daksh.employee{load_date=2024-02-13} stats: [numFiles=1, numRows=0, totalSize=822, rawDataSize=0]
OK
Time taken: 11.778 seconds
```

```
hive> select * from employee;
OK
2      Rohan      Trainee 25000      2024-02-13
1      Seon       Business Analyst      50000      2024-02-14
Time taken: 0.44 seconds, Fetched: 2 row(s)
```

```
hive> select id, name from employee order by id desc;
Query ID = root_20240207045921_ff969863-6954-40d3-9c43-c613c771f0b1
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707276530290_0002)

-----
VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED    6         6         0         0         0         0
Reducer 2 .....  SUCCEEDED    1         1         0         0         0         0
-----
VERTICES: 02/02 [=====>>] 100% ELAPSED TIME: 82.73 s
-----
OK
2      Rohan
1      Seon
Time taken: 86.367 seconds, Fetched: 2 row(s)
```

```
hive> insert into employee partition(load_date='2024-02-12') values (003,'Kartikey','HR Manager',60000);
Query ID = root_20240207050447_bb449158-aa4d-4d42-8db3-3eb13ae0ee14
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707276530290_0002)

-----
VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED    1         1         0         0         0         0
-----
VERTICES: 01/01 [=====>>] 100% ELAPSED TIME: 5.17 s
-----
Loading data to table daksh.employee partition (load_date=2024-02-12)
Partition daksh.employee{load_date=2024-02-12} stats: [numFiles=1, numRows=0, totalSize=849, rawDataSize=0]
OK
Time taken: 9.006 seconds
```

```
hive> select * from employee;
OK
3      Kartikey      HR Manager      60000      2024-02-12
2      Rohan      Trainee 25000      2024-02-13
1      Seon      Business Analyst      50000      2024-02-14
Time taken: 0.511 seconds, Fetched: 3 row(s)
```

```
hive> insert into employee partition (load_date='2024-02-11') values (004,'Jayesh','Trainee',30000);
Query ID = root_20240207051352_fd0dfe37-46ab-40b4-afe0-fb303fcd479f
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707276530290_0002)

-----
VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED    1         1         0         0         0         0
-----
VERTICES: 01/01 [=====>>] 100% ELAPSED TIME: 5.84 s
-----
Loading data to table daksh.employee partition (load_date=2024-02-11)
Partition daksh.employee{load_date=2024-02-11} stats: [numFiles=1, numRows=0, totalSize=828, rawDataSize=0]
OK
Time taken: 8.24 seconds
```

```
hive> select dept,count(*) from employee group by dept;
Query ID = root_20240207052012_e20b6321-971f-4c7f-9cee-14fe82911bb3
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707276530290_0002)

-----
VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED    12         12         0         0         7         0
Reducer 2 .....  SUCCEEDED     1          1         0         0         0         0
-----
VERTICES: 02/02 [=====>>>] 100% ELAPSED TIME: 31.03 s
-----
OK
Business Analyst      1
HR Manager            1
Trainee 2
Time taken: 32.688 seconds, Fetched: 3 row(s)
```

```
hive> update employee set name='Daksh' where id=004;
Query ID = root_20240207052217_634d1978-b562-4034-b869-12fa54b3d549
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707276530290_0002)

-----
VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED    12         12         0         0         4         0
Reducer 2 .....  SUCCEEDED     1          1         0         0         0         0
-----
VERTICES: 02/02 [=====>>>] 100% ELAPSED TIME: 39.71 s
-----
Loading data to table daksh.employee partition (load_date=null)
Time taken for load dynamic partitions : 690
Loading partition {load date=2024-02-11}
Time taken for adding to write entity : 1
Partition daksh.employee{load_date=2024-02-11} stats: [numFiles=2, numRows=0, totalSize=1670, rawDataSize=0]
OK
Time taken: 44.092 seconds
hive> select * from employee;
OK
4      Daksh   Trainee 30000   2024-02-11
3      Kartikey HR Manager 60000   2024-02-12
2      Rohan   Trainee 25000   2024-02-13
1      Seon    Business Analyst 50000   2024-02-14
Time taken: 0.177 seconds, Fetched: 4 row(s)
```



```
hive> delete from employee where name='Seon';
Query ID = root_20240207052417_99c6666f-414c-4a1f-bd20-e72f6d30e1e9
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707276530290_0002)

-----
VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED  12      12          0         0         5         0
Reducer 2 .....  SUCCEEDED   1        1          0         0         0         0
-----
VERTICES: 02/02 [=====>>>] 100% ELAPSED TIME: 38.88 s
-----
Loading data to table daksh.employee partition (load_date=null)
Time taken for load dynamic partitions : 380
Loading partition {load_date=2024-02-14}
Time taken for adding to write entity : 0
Partition daksh.employee{load_date=2024-02-14} stats: [numFiles=2, numRows=0, totalSize=1390, rawDataSize=0]
OK
Time taken: 41.935 seconds
hive> select * from employee;
OK
4      Daksh   Trainee 30000  2024-02-11
3      Kartikey      HR Manager  60000  2024-02-12
2      Rohan   Trainee 25000  2024-02-13
Time taken: 0.264 seconds, Fetched: 3 row(s)
```

```
hive> alter table employee rename to emp;
OK
Time taken: 0.871 seconds
hive> show tables;
OK
emp
values__tmp__table__1
values__tmp__table__2
values__tmp__table__3
values__tmp__table__4
values__tmp__table__5
Time taken: 0.149 seconds, Fetched: 6 row(s)
```

```
hive> alter table emp add columns(mobile int);
OK
Time taken: 0.469 seconds
hive> select * from emp;
OK
4      Daksh   Trainee 30000  NULL  2024-02-11
3      Kartikey      HR Manager  60000  NULL  2024-02-12
2      Rohan   Trainee 25000  NULL  2024-02-13
Time taken: 0.303 seconds, Fetched: 3 row(s)
```

```
hive> alter table emp change name sname String;
OK
Time taken: 0.526 seconds
```

```
hive> update emp set mobile=123 where id=002;
Query ID = root_20240207053508_efde3dcb-6179-4319-a224-6472c51343ea
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707276530290_0002)

-----
VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED    12         12         0         0         5         0
Reducer 2 .....  SUCCEEDED     1          1         0         0         0         0
-----
VERTICES: 02/02 [=====>>] 100% ELAPSED TIME: 43.17 s
-----
Loading data to table daksh.emp partition (load_date=null)
  Time taken for load dynamic partitions : 313
  Loading partition {load_date=2024-02-13}
  Time taken for adding to write entity : 19
Partition daksh.emp{load_date=2024-02-13} stats: [numFiles=2, numRows=0, totalSize=1735, rawDataSize=0]
OK
Time taken: 45.574 seconds
```

```
hive> update emp set mobile=321 where id=003;
Query ID = root_20240207053648_b9bbdeff-1303-44be-96b7-dee0f2ab51e8
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707276530290_0002)

-----
VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED    12         12         0         0         3         0
Reducer 2 .....  SUCCEEDED     1          1         0         0         0         0
-----
VERTICES: 02/02 [=====>>] 100% ELAPSED TIME: 41.20 s
-----
Loading data to table daksh.emp partition (load_date=null)
  Time taken for load dynamic partitions : 693
  Loading partition {load_date=2024-02-12}
  Time taken for adding to write entity : 0
Partition daksh.emp{load_date=2024-02-12} stats: [numFiles=2, numRows=0, totalSize=1790, rawDataSize=0]
OK
Time taken: 45.486 seconds
```

```
hive> update emp set mobile=213 where id=004;
Query ID = root_20240207053844_8ba9ed0e-b135-4885-99b6-c3effee4bc31
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707276530290_0002)

-----
VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED    12         12         0         0         4         0
Reducer 2 .....  SUCCEEDED     1          1         0         0         0         0
-----
VERTICES: 02/02 [=====>>] 100% ELAPSED TIME: 40.07 s
-----
Loading data to table daksh.emp partition (load_date=null)
  Time taken for load dynamic partitions : 598
  Loading partition {load_date=2024-02-11}
  Time taken for adding to write entity : 0
Partition daksh.emp{load_date=2024-02-11} stats: [numFiles=3, numRows=0, totalSize=2584, rawDataSize=0]
OK
Time taken: 43.797 seconds
```



```
hive> select * from emp;
OK
4      Daksh   Trainee 30000   213      2024-02-11
3      Kartikey      HR Manager 60000   321      2024-02-12
2      Rohan    Trainee 25000   123      2024-02-13
Time taken: 0.423 seconds, Fetched: 3 row(s)
```

```
hive> select sum(salary) from emp;
Query ID = root_20240207054658_3dcb9a0b-caba-4502-8e8b-dff080a76657
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707276530290_0002)

-----
VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED    12         12         0         0         5         0
Reducer 2 .....  SUCCEEDED     1          1         0         0         0         0
-----
VERTICES: 02/02 [=====>>] 100%  ELAPSED TIME: 31.49 s
-----
OK
115
Time taken: 35.563 seconds, Fetched: 1 row(s)
```

```
hive> select max(salary) from emp;
Query ID = root_20240207054903_11423223-a0af-431d-82ef-d5f101703dd5
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707276530290_0002)

-----
VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED    12         12         0         0         4         0
Reducer 2 .....  SUCCEEDED     1          1         0         0         0         0
-----
VERTICES: 02/02 [=====>>] 100%  ELAPSED TIME: 30.67 s
-----
OK
60
Time taken: 31.889 seconds, Fetched: 1 row(s)
```

```
hive> select min(salary) from emp;
Query ID = root_20240207055014_aef82bd3-61d3-4209-8738-e94061c1e9a5
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707276530290_0002)

-----
      VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED    12         12         0         0         4         0
Reducer 2 .....  SUCCEEDED     1          1         0         0         0         0
-----
VERTICES: 02/02 [=====>>] 100% ELAPSED TIME: 26.82 s
-----
OK
25
Time taken: 29.225 seconds, Fetched: 1 row(s)
```

```
hive> select avg(salary) from emp;
Query ID = root_20240207055110_43878205-006a-4813-8c54-13806ff4415d
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707276530290_0002)

-----
      VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED    12         12         0         0         7         0
Reducer 2 .....  SUCCEEDED     1          1         0         0         0         0
-----
VERTICES: 02/02 [=====>>] 100% ELAPSED TIME: 18.36 s
-----
OK
38.333333333333336
Time taken: 20.896 seconds, Fetched: 1 row(s)
```

```
hive> select count(salary) from emp;
Query ID = root_20240207055233_af0b03a7-0dfe-4cf7-9392-bb2cd18e67db
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707276530290_0002)

-----
      VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED    12         12         0         0         6         0
Reducer 2 .....  SUCCEEDED     1          1         0         0         0         0
-----
VERTICES: 02/02 [=====>>] 100% ELAPSED TIME: 27.36 s
-----
OK
3
Time taken: 28.292 seconds, Fetched: 1 row(s)
```

```
hive> select * from emp where salary>25;
OK
4      Daksh   Trainee 30      213      2024-02-11
3      Kartikey HR Manager 60      321      2024-02-12
Time taken: 0.267 seconds, Fetched: 2 row(s)
```

```
hive> select * from emp where salary<60;
OK
4      Daksh   Trainee 30      213      2024-02-11
2      Rohan   Trainee 25      123      2024-02-13
Time taken: 0.422 seconds, Fetched: 2 row(s)
```

```
hive> create table customers(id int,name String,age int, address String, salary int)
> partitioned by(load_date date)
> clustered by(id) into 3 buckets
> stored as orc tblproperties('transactional'='true');
OK
Time taken: 0.705 seconds
```

```
hive> insert into customers partition (load_date='2023-03-09') values (1,"Ross",25,"Mumbai",25000);
Query ID = root_20240207055722_d367abf3-0833-4573-95d3-fe1995a748b5
Total jobs = 1
Launching Job 1 out of 1
```

Status: Running (Executing on YARN cluster with App id application\_1707276530290\_0002)

```
-----
VERTICES    STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 ..... SUCCEEDED    1          1          0          0          0          0
-----
VERTICES: 01/01 [=====>>] 100% ELAPSED TIME: 4.87 s
-----
```

```
Loading data to table daksh.customers partition (load_date=2023-03-09)
Partition daksh.customers{load_date=2023-03-09} stats: [numFiles=1, numRows=0, totalSize=865, rawDataSize=0]
OK
Time taken: 7.301 seconds
```

```
hive> insert into customers partition (load_date='2023-03-09') values (2,"Mike",27,"Bhopal",35000);
Query ID = root_20240207055836_f65726a4-5d1a-4a60-a647-5146f7f40052
Total jobs = 1
Launching Job 1 out of 1
```

Status: Running (Executing on YARN cluster with App id application\_1707276530290\_0002)

```
-----
VERTICES    STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 ..... SUCCEEDED    1          1          0          0          0          0
-----
VERTICES: 01/01 [=====>>] 100% ELAPSED TIME: 3.25 s
-----
```

```
Loading data to table daksh.customers partition (load_date=2023-03-09)
Partition daksh.customers{load_date=2023-03-09} stats: [numFiles=2, numRows=0, totalSize=1734, rawDataSize=0]
OK
Time taken: 5.806 seconds
```

```
hive> insert into customers partition(load_date='2023-03-10') values (3,"Albin",24,"Pune",50000);
Query ID = root_20240207055932_685c48e4-2fef-493b-9712-ab7ffca11b39
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707276530290_0002)

-----
VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED      1          1           0           0           0           0
-----
VERTICES: 01/01 [=====>>] 100% ELAPSED TIME: 4.05 s
-----
Loading data to table daksh.customers partition (load_date=2023-03-10)
Partition daksh.customers{load_date=2023-03-10} stats: [numFiles=1, numRows=0, totalSize=864, rawDataSize=0]
OK
Time taken: 6.615 seconds
```

```
hive> select * from customers;
OK
1      Ross      25      Mumbai  25000    2023-03-09
2      Mike      27      Bhopal   35000    2023-03-09
3      Albin     24      Pune     50000    2023-03-10
Time taken: 0.297 seconds, Fetched: 3 row(s)
```

```
hive> create table orders(oid int,customer_id int,amount int);
OK
Time taken: 0.566 seconds
```

```
hive> insert into orders values(101,2,20000);
Query ID = root_20240207060151_51dc8f15-93c8-41b1-87c0-de7ea9ed1482
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707276530290_0002)

-----
VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED      1          1           0           0           0           0
-----
VERTICES: 01/01 [=====>>] 100% ELAPSED TIME: 3.96 s
-----
Loading data to table daksh.orders
Table daksh.orders stats: [numFiles=1, numRows=1, totalSize=12, rawDataSize=11]
OK
Time taken: 6.773 seconds
```

```
hive> insert into orders values(102,2,25000);
Query ID = root_20240207060214_b4ff387f-6e25-423b-8ae0-cb5b2d53666b
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707276530290_0002)

-----
      VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED      1          1          0          0          0          0
-----
VERTICES: 01/01 [=====>>] 100%  ELAPSED TIME: 0.82 s
-----
Loading data to table daksh.orders
Table daksh.orders stats: [numFiles=2, numRows=2, totalSize=24, rawDataSize=22]
OK
Time taken: 1.642 seconds
```

```
hive> insert into orders values(104,4,10000);
Query ID = root_20240207060233_2b8737fd-3a08-4328-aad7-d507a4c83b3d
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707276530290_0002)

-----
      VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED      1          1          0          0          0          0
-----
VERTICES: 01/01 [=====>>] 100%  ELAPSED TIME: 3.92 s
-----
Loading data to table daksh.orders
Table daksh.orders stats: [numFiles=3, numRows=3, totalSize=36, rawDataSize=33]
OK
Time taken: 4.619 seconds
```

```
hive> insert into orders values (103,1,30000);
Query ID = root_20240207060416_f2f7d175-633d-4b17-8a34-0da6cd48c881
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707276530290_0002)

-----
      VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED      1          1          0          0          0          0
-----
VERTICES: 01/01 [=====>>] 100%  ELAPSED TIME: 3.06 s
-----
Loading data to table daksh.orders
Table daksh.orders stats: [numFiles=4, numRows=4, totalSize=48, rawDataSize=44]
OK
Time taken: 4.322 seconds
```

```
hive> select * from orders;
OK
101      2      20000
102      2      25000
104      4      10000
103      1      30000
Time taken: 0.106 seconds, Fetched: 4 row(s)
```

```
hive> select * from customers c join orders o
  > on (c.id=o.customer_id);
Query ID = root_20240207060527_e59255ab-8cd0-4be7-9d5a-71d6717f2bc0
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707276530290_0002)

-----
      VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED      6         6         0         0         0         0
Map 2 .....  SUCCEEDED      1         1         0         0         0         0
-----
VERTICES: 02/02 [=====>>] 100%  ELAPSED TIME: 23.79 s
-----
OK
1      Ross      25      Mumbai  25000    2023-03-09      103      1      30000
2      Mike      27      Bhopal   35000    2023-03-09      101      2      20000
2      Mike      27      Bhopal   35000    2023-03-09      102      2      25000
Time taken: 24.925 seconds, Fetched: 3 row(s)
```

```
hive> select * from customers c left outer join orders o
  > on (c.id=o.customer_id);
Query ID = root_20240207060641_01020698-1f1a-465a-9c3b-a39b7d9e910b
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707276530290_0002)

-----
      VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED      6         6         0         0         0         0
Map 2 .....  SUCCEEDED      1         1         0         0         0         0
-----
VERTICES: 02/02 [=====>>] 100%  ELAPSED TIME: 31.89 s
-----
OK
1      Ross      25      Mumbai  25000    2023-03-09      103      1      30000
2      Mike      27      Bhopal   35000    2023-03-09      101      2      20000
2      Mike      27      Bhopal   35000    2023-03-09      102      2      25000
3      Albin     24      Pune     50000    2023-03-10      NULL     NULL     NULL
Time taken: 33.57 seconds, Fetched: 4 row(s)
```

```
hive> select * from customers c right outer join orders o
  > on (c.id=o.customer_id);
Query ID = root_20240207060756_3f15f394-0f3e-4715-94e4-3dbfdaee97ce
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707276530290_0002)

-----
      VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED      6          6          0          0          0          0
Map 2 .....  SUCCEEDED      1          1          0          0          0          0
-----
VERTICES: 02/02 [=====>>>] 100%  ELAPSED TIME: 30.18 s
-----
OK
2      Mike      27      Bhopal      35000      2023-03-09      101      2      20000
2      Mike      27      Bhopal      35000      2023-03-09      102      2      25000
NULL   NULL      NULL      NULL      NULL      NULL      104      4      10000
1      Ross      25      Mumbai      25000      2023-03-09      103      1      30000
Time taken: 31.388 seconds, Fetched: 4 row(s)
```

```
hive> select * from customers c full outer join orders o on (c.id=o.customer_id);
Query ID = root_20240207060915_167ea983-040a-4e6c-922b-9f222dc0913c
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707276530290_0002)

-----
      VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED      6          6          0          0          0          0
Map 3 .....  SUCCEEDED      1          1          0          0          0          0
Reducer 2 .....  SUCCEEDED      1          1          0          0          0          0
-----
VERTICES: 03/03 [=====>>>] 100%  ELAPSED TIME: 50.69 s
-----
OK
1      Ross      25      Mumbai      25000      2023-03-09      103      1      30000
2      Mike      27      Bhopal      35000      2023-03-09      101      2      20000
2      Mike      27      Bhopal      35000      2023-03-09      102      2      25000
3      Albin      24      Pune      50000      2023-03-10      NULL      NULL      NULL
NULL   NULL      NULL      NULL      NULL      NULL      104      4      10000
Time taken: 52.647 seconds, Fetched: 5 row(s)
```

```
hive> create view if not exists customers_vw as select * from customers where address="Mumbai";
OK
Time taken: 0.889 seconds
```

```
hive> select * from customers_vw;
OK
1      Ross      25      Mumbai      25000      2023-03-09
Time taken: 0.588 seconds, Fetched: 1 row(s)
```

```
hive> alter view customers_vw as select * from customers;
OK
Time taken: 0.313 seconds
```



```
hive> select * from customers_vw;  
OK  
1      Ross      25      Mumbai  25000   2023-03-09  
2      Mike      27      Bhopal  35000   2023-03-09  
3      Albin     24      Pune    50000   2023-03-10  
Time taken: 0.15 seconds, Fetched: 3 row(s)
```

```
hive> drop view if exists customers_vw;  
OK  
Time taken: 0.925 seconds
```