

Name: Purna Sunil Jadhav

Sap ID: 60004220127

Batch: C22

Course: Big Data Infrastructure Laboratory

Course Code: DT19CEEL6011

EXPERIMENT 04

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

THEORY:

- Hive is a data warehouse infrastructure tool to process the structured data in Hadoop.
- It resides on top of Hadoop to summarize Big Data, and makes querying and analyzing easy.
- Initially Hive was developed by Facebook, later the Apache Software Foundation took it up and developed it further as an open source under the name Apache Hive.
- It is used by different companies. For example, Amazon uses it in Amazon Elastic MapReduce
- Features of Hive:
 - It stores the schema in database & processed data into HDFS
 - It is designed for OLAP
 - It is familiar, fast, scalable & extensible.

COMMANDS: [root@sandbox ~]# live

- 1) show databases;
→ shows databases already existing in hive.
- 2) create database student;
→ creates a new database, here, student.
- 3) use student;
→ using the database for executing queries.
- 4) create table student (id int, name varchar(20),
branch varchar(20), mobile int)
PARTITIONED by (load_date date)
CLUSTERED by (id) INTO 3 BUCKETS
STORED as ORC TBLPROPERTIES ('transactional'
→ creates table in the DB. = 'true');
- 5) insert into student partition (load_date =
'2023-03-09') values (101, 'Prema', 'IT', 12346);
- 6) select * from student;
→ To display contents of table.
- 7) select id, name from student order by id desc;
→ order by clause (ordering the result
in descending order; by default
ascending order).

- 8) `select branch, count (*) from student
group by branch;`
→ Grouping the result with respect to branch here.
- 9) `update student set name = 'Diksha' where
id = 101;`
→ updates the value in the table
- 10) `delete from student where name = 'Diksha';`
→ Delete the entry from the table.
- 11) `alter table student rename to stud;`
→ renaming a table
- 12) `alter table stud add columns (cgpa double);`
→ add new column in the table
- 13) `alter table stud change name sname string;`
→ Changing the column name with alter statement, also changing the the datatype from varchar to string.
- 14) `select sum(cgpa), max(cgpa), min(cgpa),
avg(cgpa), count(cgpa) from stud;`
→ Performing aggregate functions (sum, max, min, avg, count) on the table.

15) Relational operations in Hive

```
select * from stud where cgpa < 9.0;
```

```
select * from stud where cgpa >= 7.5;
```

16) Joins in Hive

→ Inner:

```
select * from customers c join orders  
o on (c.id = o.customer_id);
```

→ full outer:

```
select * from customers c full  
outer join order o on (c.id = o.cust-id);
```

17) Views in Hive

```
create view if not exists customers_vw  
as select * from customers where  
address = "Mumbai";
```

→ creating view

18) Dropping the view

```
drop view if exists customers_vw;
```

CONCLUSION:

we have successfully executed HIVE queries using HQL in Hadoop.



Academic Year: 2022-2023

Name:	Prerna Sunil Jadhav
Sap Id:	60004220127
Class:	T. Y. B. Tech (Computer Engineering)
Course:	Big Data Infrastructure Laboratory
Course Code:	DJ19CEEL6011
Experiment No.:	04

AIM: Execute Hive commands to load, insert, update or delete data in tables.

OUTPUT:

```
sandbox login: root
root@sandbox.hortonworks.com's password:
Last login: Wed Feb  7 03:26:53 2024 from 172.17.0.2
[root@sandbox ~]# hive

Logging initialized using configuration in file:/etc/hive/2.5.0.0-1245/0/hive-log4j.properties
hive> show databases;
OK
default
foodmart
xademo
Time taken: 2.348 seconds, Fetched: 3 row(s)
hive> create database prernaDB;
OK
Time taken: 2.276 seconds
hive> use prernaDB;
OK
Time taken: 0.274 seconds
hive> create table prerna(id int, name varchar(20), branch varchar(20), mobile int)
> PARTITIONED BY (load_date date)
> CLUSTERED BY (id) INTO 3 BUCKETS
> STORED AS ORC TBLPROPERTIES ('TRANSACTIONAL'='true');
OK
Time taken: 1.084 seconds
hive> describe prerna;
OK
id                int
name              varchar(20)
branch            varchar(20)
mobile            int
load_date         date

# Partition Information
# col_name        data_type        comment
load_date         date
Time taken: 0.604 seconds, Fetched: 10 row(s)
hive> describe prernaDB DEFAULT;
FAILED: SemanticException [Error 10001]: Table not found prernaDB
hive> describe DATABASE prernaDB;
OK
prernadb          hdfs://sandbox.hortonworks.com:8020/apps/hive/warehouse/prernadb.db    roo
SER
Time taken: 0.308 seconds, Fetched: 1 row(s)
hive> insert into prerna partition (load_date = '01-01-2024') values (101,'Prerna','Computer',98765
);
FAILED: SemanticException [Error 10248]: Cannot add partition column load_date of type string as it
ot be converted to type date
hive> insert into prerna partition (load_date = 01-01-2024) values (101,'Prerna','Computer',9876543
```



Academic Year: 2022-2023

```
hive> insert into prerna partition (load_date = '2024-02-14') values (101,'Prerna','Computer',9876543210);
Query ID = root_20240214033300_12570c8c-8e3b-4b44-af8a-4168fbaf705d
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_1707880073946_0002)

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

VERTICES: 01/01 [=====>>>] 100% ELAPSED TIME: 9.16 s

```
Loading data to table prernadb.prerna partition (load_date=2024-02-14)
Partition prernadb.prerna{load_date=2024-02-14} stats: [numFiles=1, numRows=0, totalSize=839, rawDataSize=0]
OK
Time taken: 23.035 seconds
```

```
hive> insert into prerna partition (load_date = '2024-02-14') values (102,'Akshata','EXTC',999888777666);
Query ID = root_20240214033428_a06736de-b8a0-47cb-8721-827ff194d39e
Total jobs = 1
Launching Job 1 out of 1
```

Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)

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

VERTICES: 01/01 [=====>>>] 100% ELAPSED TIME: 5.15 s

```
Loading data to table prernadb.prerna partition (load_date=2024-02-14)
Partition prernadb.prerna{load_date=2024-02-14} stats: [numFiles=2, numRows=0, totalSize=1669, rawDataSize=0]
OK
Time taken: 8.092 seconds
```

```
hive> insert into prerna partition (load_date = '2024-02-14') values (103,'Diksha','Mechanical',81639852);
Query ID = root_20240214033523_59c07c0e-94e6-441c-a9e7-bd246c9870e5
Total jobs = 1
Launching Job 1 out of 1
```




Academic Year: 2022-2023

```
-----  
VERTICES    STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED  
-----  
Map 1 ..... SUCCEEDED    1         1         0         0         0         0  
-----  
VERTICES: 01/01 [=====>>] 100% ELAPSED TIME: 3.50 s  
-----  
Loading data to table prernadb.prerna partition (load_date=2024-02-14)  
Partition prernadb.prerna{load_date=2024-02-14} stats: [numFiles=3, numRows=0, totalSize=2545, rawDataSize=0]  
OK  
Time taken: 5.912 seconds  
hive> select * from prerna;  
OK  
101      Prerna Computer      NULL      2024-02-14  
102      Akshata EXTC      NULL      2024-02-14  
103      Diksha Mechanical      81639852      2024-02-14  
Time taken: 0.487 seconds, Fetched: 3 row(s)  
hive> select id, name from prerna order by id desc;  
Query ID = root_20240214033840_1c1aee1-09fb-48f6-9c1a-585d9756a3b6  
Total jobs = 1  
Launching Job 1 out of 1  
  
Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)  
  
-----  
VERTICES    STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED  
-----  
Map 1 ..... SUCCEEDED    3         3         0         0         0         0  
Reducer 2 ..... SUCCEEDED    1         1         0         0         0         0  
-----  
VERTICES: 02/02 [=====>>] 100% ELAPSED TIME: 28.00 s  
-----  
OK  
103      Diksha  
102      Akshata  
101      Prerna  
Time taken: 29.595 seconds, Fetched: 3 row(s)  
hive> insert into prerna partition (load_date = '2024-02-14') values (104,'Yash','EXTC',816392);  
Query ID = root_20240214034312_87ad898b-9224-4e40-b42a-a1a0ea30b868  
Total jobs = 1  
Launching Job 1 out of 1  
  
Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)
```



Academic Year: 2022-2023

```
-----
VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED    1         1         0         0         0         0
-----
VERTICES: 01/01 [=====>>>] 100% ELAPSED TIME: 6.61 s
-----
Loading data to table prernadb.prerna partition (load_date=2024-02-14)
Partition prernadb.prerna{load_date=2024-02-14} stats: [numFiles=4, numRows=0, totalSize=3364, rawD
ataSize=0]
OK
Time taken: 9.346 seconds
hive> insert into prerna partition (load_date = '2024-02-14') values (105,'Krishna','Mechanical',85
852);
Query ID = root_20240214034353_07eb7d27-d08a-44d6-b007-25a25e9a6609
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)

-----
VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED    1         1         0         0         0         0
-----
VERTICES: 01/01 [=====>>>] 100% ELAPSED TIME: 2.90 s
-----
Loading data to table prernadb.prerna partition (load_date=2024-02-14)
Partition prernadb.prerna{load_date=2024-02-14} stats: [numFiles=5, numRows=0, totalSize=4236, rawD
ataSize=0]
OK
Time taken: 5.408 seconds
hive> insert into prerna partition (load_date = '2024-02-14') values (106,'Kajal','Chemical',84555)
;
Query ID = root_20240214034438_bd92ae2e-dbbf-40f5-b064-69ea0dc9613c
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)

-----
VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED    1         1         0         0         0         0
-----
VERTICES: 01/01 [=====>>>] 100% ELAPSED TIME: 2.25 s
-----
```




Academic Year: 2022-2023

```
-----
Loading data to table prernadb.prerna partition (load_date=2024-02-14)
Partition prernadb.prerna{load_date=2024-02-14} stats: [numFiles=6, numRows=0, totalSize=5095, rawD
ataSize=0]
OK
Time taken: 4.007 seconds
hive> select * from prerna;
OK
101    Prerna    Computer    NULL    2024-02-14
102    Akshata  EXTC      NULL    2024-02-14
103    Diksha    Mechanical 81639852 2024-02-14
104    Yash      EXTC      816392  2024-02-14
105    Krishna  Mechanical 85852   2024-02-14
106    Kajal    Chemical  84555   2024-02-14
Time taken: 0.166 seconds, Fetched: 6 row(s)
hive> SELECT branch, count(*) FROM prerna GROUP BY branch;
Query ID = root_20240214034613_bd0b8fd4-fbfa-461a-842e-e02c1e834dca
Total jobs = 1
Launching Job 1 out of 1
```

Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)

	VERTICES	STATUS	TOTAL	COMPLETED	RUNNING	PENDING	FAILED	KILLED
Map 1		SUCCEEDED	3	3	0	0	0	0
Reducer 2		SUCCEEDED	1	1	0	0	0	0

VERTICES: 02/02 [=====>>] 100% ELAPSED TIME: 12.66 s

```
-----
OK
Chemical      1
Computer      1
EXTC          2
Mechanical    2
Time taken: 14.684 seconds, Fetched: 4 row(s)
hive> update prerna set mobile=888552 where id=101;
Query ID = root_20240214035000_29c24cba-49cd-47b3-8a8e-f03d592c0bfd
Total jobs = 1
Launching Job 1 out of 1
```

Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)

	VERTICES	STATUS	TOTAL	COMPLETED	RUNNING	PENDING	FAILED	KILLED
Map 1		SUCCEEDED	3	3	0	0	0	0



Academic Year: 2022-2023

```
Partition prernadb.prerna{load_date=2024-02-14} stats: [numFiles=7, numRows=0, totalSize=5965, rawDataSize=0]
OK
Time taken: 22.841 seconds
hive> update prerna set mobile=888552 where id=102;
Query ID = root_20240214035201_1260ea37-75c3-419e-8f0e-8957f8d430f4
Total jobs = 1
Launching Job 1 out of 1
```

Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)

VERTICES	STATUS	TOTAL	COMPLETED	RUNNING	PENDING	FAILED	KILLED
Map 1	SUCCEEDED	3	3	0	0	0	0
Reducer 2	SUCCEEDED	1	1	0	0	0	0

VERTICES: 02/02 [=====>>>] 100% ELAPSED TIME: 11.20 s

Loading data to table prernadb.prerna partition (load_date=null)

Time taken for load dynamic partitions : 699

Loading partition {load_date=2024-02-14}

Time taken for adding to write entity : 0

```
Partition prernadb.prerna{load_date=2024-02-14} stats: [numFiles=8, numRows=0, totalSize=6836, rawDataSize=0]
OK
```

Time taken: 14.194 seconds

```
hive> select * from prerna;
```

OK

101	Prerna	Computer	888552	2024-02-14
102	Akshata	EXTC	888552	2024-02-14
103	Diksha	Mechanical	81639852	2024-02-14
104	Yash	EXTC	816392	2024-02-14
105	Krishna	Mechanical	85852	2024-02-14
106	Kajal	Chemical	84555	2024-02-14

Time taken: 0.187 seconds, Fetched: 6 row(s)

```
hive> delete from prerna where name='Prerna';
```

Query ID = root_20240214035417_f936f3cd-97b7-4415-9784-769bff40fe42

Total jobs = 1

Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)

VERTICES	STATUS	TOTAL	COMPLETED	RUNNING	PENDING	FAILED	KILLED
Map 1	SUCCEEDED	3	3	0	0	0	0
Reducer 3	SUCCEEDED	1	1	0	0	0	0



Academic Year: 2022-2023

```
hive> select * from prerna;
OK
102    Akshata EXTC      888552  2024-02-14
103    Diksha  Mechanical 81639852 2024-02-14
104    Yash   EXTC      816392  2024-02-14
105    Krishna Mechanical 85852   2024-02-14
106    Kajal  Chemical   84555   2024-02-14
Time taken: 0.163 seconds, Fetched: 5 row(s)
hive> alter table prerna rename to prernaTable
> ;
OK
Time taken: 0.574 seconds
hive> show tables;
OK
prernatable
values_tmp_table_1
values_tmp_table_2
values_tmp_table_3
values_tmp_table_4
values_tmp_table_5
values_tmp_table_6
values_tmp_table_7
values_tmp_table_8
values_tmp_table_9
Time taken: 0.16 seconds, Fetched: 10 row(s)
hive> alter table prernaTable add columns(cgpa double);
OK
Time taken: 0.294 seconds
hive> describe prernaTable;
OK
id                int
name              varchar(20)
branch            varchar(20)
mobile            int
cgpa              double
load_date         date

# Partition Information
# col_name        data_type        comment
load_date         date
Time taken: 0.15 seconds, Fetched: 11 row(s)
hive> select * from prernaTable;
OK
102    Akshata EXTC      888552  NULL    2024-02-14
103    Diksha  Mechanical 81639852 NULL    2024-02-14
104    Yash   EXTC      816392  NULL    2024-02-14
105    Krishna Mechanical 85852   NULL    2024-02-14
106    Kajal  Chemical   84555   NULL    2024-02-14
Time taken: 0.136 seconds, Fetched: 5 row(s)
```



Academic Year: 2022-2023

```
hive> alter table prernaTable CHANGE name sname String;
OK
Time taken: 0.351 seconds
hive> update prernaTable set cgpa=9.3 where id=102;
Query ID = root_20240214035910_5623fb72-56df-46d1-81b8-40abe35bca95
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)

-----
VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED    3          3          0          0          0          0
Reducer 2 .....  SUCCEEDED    1          1          0          0          0          0
-----
VERTICES: 02/02 [=====>>>] 100% ELAPSED TIME: 18.41 s
-----
Loading data to table prernadb.prernatable partition (load_date=null)
  Time taken for load dynamic partitions : 249
  Loading partition {load_date=2024-02-14}
  Time taken for adding to write entity : 0
Partition prernadb.prernatable[load_date=2024-02-14] stats: [numFiles=10, numRows=0, totalSize=8327, rawDataSize=0]
OK
Time taken: 21.172 seconds
hive> select * from prernaTable;
OK
102  Akshata  EXTC    888552  9.3    2024-02-14
103  Diksha   Mechanical  81639852  NULL   2024-02-14
104  Yash     EXTC    816392  NULL   2024-02-14
105  Krishna  Mechanical  85852    NULL   2024-02-14
106  Kajal    Chemical  84555    NULL   2024-02-14
Time taken: 0.603 seconds, Fetched: 5 row(s)
hive> select sum(cgpa) from prernaTable;
Query ID = root_20240214040100_4181bffb-a65b-46ab-be9a-38edd994780d
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)

-----
VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED    3          3          0          0          0          0
Reducer 2 .....  SUCCEEDED    1          1          0          0          0          0
-----
```




Academic Year: 2022-2023

```
hive> alter table prernaTable CHANGE name sname String;
OK
Time taken: 0.351 seconds
hive> update prernaTable set cgpa=9.3 where id=102;
Query ID = root_20240214035910_5623fb72-56df-46d1-81b8-40abe35bca95
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)

-----
VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED    3          3          0          0          0          0
Reducer 2 .....  SUCCEEDED    1          1          0          0          0          0
-----
VERTICES: 02/02 [=====>>>] 100% ELAPSED TIME: 18.41 s
-----
Loading data to table prernadb.prernatable partition (load_date=null)
  Time taken for load dynamic partitions : 249
  Loading partition {load_date=2024-02-14}
  Time taken for adding to write entity : 0
Partition prernadb.prernatable[load_date=2024-02-14] stats: [numFiles=10, numRows=0, totalSize=8327, rawDataSize=0]
OK
Time taken: 21.172 seconds
hive> select * from prernaTable;
OK
102  Akshata  EXTC    888552  9.3    2024-02-14
103  Diksha   Mechanical  81639852  NULL   2024-02-14
104  Yash     EXTC    816392  NULL   2024-02-14
105  Krishna  Mechanical  85852    NULL   2024-02-14
106  Kajal    Chemical  84555    NULL   2024-02-14
Time taken: 0.603 seconds, Fetched: 5 row(s)
hive> select sum(cgpa) from prernaTable;
Query ID = root_20240214040100_4181bffb-a65b-46ab-be9a-38edd994780d
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)

-----
VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED    3          3          0          0          0          0
Reducer 2 .....  SUCCEEDED    1          1          0          0          0          0
-----
```



Academic Year: 2022-2023

```
hive> alter table prernaTable CHANGE name sname String;
OK
Time taken: 0.351 seconds
hive> update prernaTable set cgpa=9.3 where id=102;
Query ID = root_20240214035910_5623fb72-56df-46d1-81b8-40abe35bca95
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)

-----
VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED    3          3          0          0          0          0
Reducer 2 ..... SUCCEEDED    1          1          0          0          0          0
-----
VERTICES: 02/02 [=====>>>] 100% ELAPSED TIME: 18.41 s
-----
Loading data to table prernadb.prernatable partition (load_date=null)
  Time taken for load dynamic partitions : 249
  Loading partition {load_date=2024-02-14}
  Time taken for adding to write entity : 0
Partition prernadb.prernatable[load_date=2024-02-14] stats: [numFiles=10, numRows=0, totalSize=8327, rawDataSize=0]
OK
Time taken: 21.172 seconds
hive> select * from prernaTable;
OK
102  Akshata  EXTC    888552  9.3    2024-02-14
103  Diksha   Mechanical  81639852  NULL   2024-02-14
104  Yash     EXTC    816392  NULL   2024-02-14
105  Krishna  Mechanical  85852    NULL   2024-02-14
106  Kajal    Chemical  84555    NULL   2024-02-14
Time taken: 0.603 seconds, Fetched: 5 row(s)
hive> select sum(cgpa) from prernaTable;
Query ID = root_20240214040100_4181bffb-a65b-46ab-be9a-38edd994780d
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)

-----
VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED    3          3          0          0          0          0
Reducer 2 ..... SUCCEEDED    1          1          0          0          0          0
-----
```




Academic Year: 2022-2023

```
hive> update prernaTable set cgpa=8.4 where id=103;
Query ID = root_20240214040138_ab9bc15a-55bf-4aa6-a8e1-1324aa93e057
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)

-----
      VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED      3          3          0          0          0          0
Reducer 2 .....  SUCCEEDED      1          1          0          0          0          0
-----
VERTICES: 02/02 [=====>>>] 100% ELAPSED TIME: 8.70 s
-----
Loading data to table prernadb.prernatable partition (load_date=null)
  Time taken for load dynamic partitions : 232
  Loading partition {load_date=2024-02-14}
  Time taken for adding to write entity : 0
Partition prernadb.prernatable{load_date=2024-02-14} stats: [numFiles=11, numRows=0, totalSize=9315,
, rawDataSize=0]
OK
Time taken: 10.29 seconds
hive> update prernaTable set cgpa=9.65 where id=106;
Query ID = root_20240214040217_575fc595-0b96-4131-9562-fee04bd1541c
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)

-----
      VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED      3          3          0          0          0          0
Reducer 2 .....  SUCCEEDED      1          1          0          0          0          0
-----
VERTICES: 02/02 [=====>>>] 100% ELAPSED TIME: 9.02 s
-----
Loading data to table prernadb.prernatable partition (load_date=null)
  Time taken for load dynamic partitions : 164
  Loading partition {load_date=2024-02-14}
  Time taken for adding to write entity : 0
Partition prernadb.prernatable{load_date=2024-02-14} stats: [numFiles=12, numRows=0, totalSize=10288,
, rawDataSize=0]
OK
Time taken: 11.294 seconds
```



Academic Year: 2022-2023

```
hive> select sum(cgpa) from prernaTable;
Query ID = root_20240214040233_0a24b166-1f88-44d2-8eca-b7110296dc79
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)

-----
VERTICES    STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 ..... SUCCEEDED    3          3          0          0          2          0
Reducer 2 ..... SUCCEEDED    1          1          0          0          0          0
-----
VERTICES: 02/02 [=====>>] 100% ELAPSED TIME: 8.86 s
-----
OK
27.35
Time taken: 9.513 seconds, Fetched: 1 row(s)
hive> select max(cgpa),min(cgpa), avg(cgpa) from prernaTable;
Query ID = root_20240214040308_b65aae0a-1aae-4212-9630-885ea4e90d3e
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)

-----
VERTICES    STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 ..... SUCCEEDED    3          3          0          0          0          0
Reducer 2 ..... SUCCEEDED    1          1          0          0          0          0
-----
VERTICES: 02/02 [=====>>] 100% ELAPSED TIME: 6.78 s
-----
OK
9.65    8.4    9.116666666666667
Time taken: 7.665 seconds, Fetched: 1 row(s)
hive> select * from prernaTable where cgpa<9;
OK
103    Diksha    Mechanical    81639852    8.4    2024-02-14
Time taken: 0.4 seconds, Fetched: 1 row(s)
hive> select * from prernaTable where cgpa>9;
OK
102    Akshata    EXTC    888552    9.3    2024-02-14
106    Kajal    Chemical    84555    9.65    2024-02-14
Time taken: 0.256 seconds, Fetched: 2 row(s)
```

```
hive> select * from prernaTable where cgpa>9;
OK
102    Akshata    EXTC    888552    9.3    2024-02-14
106    Kajal    Chemical    84555    9.65    2024-02-14
Time taken: 0.256 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: 0.418 seconds
```




Academic Year: 2022-2023

```
hive> insert into customers partition (load_date = '2024-02-14') values (1,"Ross",25,"Mumbai",25000
);
Query ID = root_20240214041011_538486b4-221d-4a0c-9ac0-e5cc95c6b26e
Total jobs = 1
Launching Job 1 out of 1
```

Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)

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

VERTICES: 01/01 [=====>>>] 100% ELAPSED TIME: 4.01 s

Loading data to table prernadb.customers partition (load_date=2024-02-14)
Partition prernadb.customers{load_date=2024-02-14} stats: [numFiles=1, numRows=0, totalSize=864, rawDataSize=0]
OK

Time taken: 8.401 seconds

```
hive> insert into customers partition (load_date = '2024-02-14') values (2,"Mike",27,"Bhopal",35000
);
Query ID = root_20240214041109_7c8069e3-274d-4682-b846-d66574964db2
Total jobs = 1
Launching Job 1 out of 1
```

Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)

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

VERTICES: 01/01 [=====>>>] 100% ELAPSED TIME: 2.95 s

Loading data to table prernadb.customers partition (load_date=2024-02-14)
Partition prernadb.customers{load_date=2024-02-14} stats: [numFiles=2, numRows=0, totalSize=1735, rawDataSize=0]
OK

Time taken: 4.776 seconds

```
hive> insert into customers partition (load_date = '2024-02-14') values(3,"Albin",24,"Pune",50000);
```

```
Query ID = root_20240214041218_76da6ad6-be77-4eed-8cc1-ad6d89100add
Total jobs = 1
Launching Job 1 out of 1
```



Academic Year: 2022-2023

```
Loading data to table prernadb.customers partition (load_date=2024-02-14)
Partition prernadb.customers{load_date=2024-02-14} stats: [numFiles=3, numRows=0, totalSize=2599, r
awDataSize=0]
OK
Time taken: 5.092 seconds
hive> select * from customers;
OK
1      Ross    25      Mumbai  25000   2024-02-14
2      Mike    27      Bhopal   35000   2024-02-14
3      Albin   24      Pune     50000   2024-02-14
Time taken: 0.207 seconds, Fetched: 3 row(s)
hive> create table orders(oid int, customer_id int, amount int);
OK
Time taken: 0.305 seconds
hive> insert into orders values(101, 2, 20000);
Query ID = root_20240214041336_181bc368-0c3d-478a-8f48-6e5fc61b814f
Total jobs = 1
Launching Job 1 out of 1
```

Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)

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

VERTICES: 01/01 [=====>>>] 100% ELAPSED TIME: 3.13 s

```
Loading data to table prernadb.orders
Table prernadb.orders stats: [numFiles=1, numRows=1, totalSize=12, rawDataSize=11]
OK
Time taken: 3.995 seconds
hive> insert into orders values(102, 2, 25000);
Query ID = root_20240214041344_bbf57554-0785-4cf2-a53d-f63f272048fc
Total jobs = 1
Launching Job 1 out of 1
```

Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)

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

VERTICES: 01/01 [=====>>>] 100% ELAPSED TIME: 0.41 s



Academic Year: 2022-2023

```
Loading data to table prernadb.orders
Table prernadb.orders stats: [numFiles=2, numRows=2, totalSize=24, rawDataSize=22]
OK
Time taken: 1.153 seconds
hive> insert into orders values(104,4,10000);
Query ID = root_20240214041358_37fc63e6-e6ad-4791-9fc7-da25306e28c7
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)

-----
      VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED      1          1          0          0          0          0
-----
VERTICES: 01/01 [=====>>] 100% ELAPSED TIME: 0.40 s
-----

Loading data to table prernadb.orders
Table prernadb.orders stats: [numFiles=3, numRows=3, totalSize=36, rawDataSize=33]
OK
Time taken: 1.545 seconds
hive> select * from orders;
OK
101      2      20000
102      2      25000
104      4      10000
Time taken: 0.149 seconds, Fetched: 3 row(s)
hive> select * from customers c join orders o on (c.id = o.customer_id);
Query ID = root_20240214041503_be050749-b1aa-463a-944c-436f82383c07
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)

-----
      VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED      3          3          0          0          0          0
Map 2 .....  SUCCEEDED      1          1          0          0          0          0
-----
VERTICES: 02/02 [=====>>] 100% ELAPSED TIME: 9.06 s
-----

OK
2      Mike      27      Bhopal      35000      2024-02-14      101      2      20000
2      Mike      27      Bhopal      35000      2024-02-14      102      2      25000
Time taken: 10.074 seconds, Fetched: 2 row(s)
```



Academic Year: 2022-2023

```
hive> select * from customers c left outer join orders o on (c.id = o.customer_id);
Query ID = root_20240214041536_d89dfa9f-9462-4923-8d02-7d8e8b8ab3a4
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)

-----
      VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED      3          3          0          0          0          0
Map 2 .....  SUCCEEDED      1          1          0          0          0          0
-----
VERTICES: 02/02 [=====] 100% ELAPSED TIME: 6.58 s
-----
OK
1      Ross    25      Mumbai  25000  2024-02-14      NULL      NULL      NULL
3      Albin   24      Pune    50000  2024-02-14      NULL      NULL      NULL
2      Mike    27      Bhopal  35000  2024-02-14      101       2        20000
2      Mike    27      Bhopal  35000  2024-02-14      102       2        25000
Time taken: 7.505 seconds, Fetched: 4 row(s)
hive> select * from customers c right outer join orders o on (c.id = o.customer_id);
Query ID = root_20240214041618_8eda5f4f-db8b-4109-b53b-da7d25e646e4
Total jobs = 1
Launching Job 1 out of 1

Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)

-----
      VERTICES      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 .....  SUCCEEDED      3          3          0          0          0          0
Map 2 .....  SUCCEEDED      1          1          0          0          0          0
-----
VERTICES: 02/02 [=====] 100% ELAPSED TIME: 6.68 s
-----
OK
2      Mike    27      Bhopal  35000  2024-02-14      101       2        20000
2      Mike    27      Bhopal  35000  2024-02-14      102       2        25000
NULL   NULL     NULL     NULL     NULL    104       4        10000
Time taken: 9.036 seconds, Fetched: 3 row(s)
hive> select * from customers c full outer join orders o on (c.id = o.customer_id);
Query ID = root_20240214041637_8c91b21b-62d9-47f2-89de-66cd979517f3
Total jobs = 1
Launching Job 1 out of 1
```



Academic Year: 2022-2023

Status: Running (Executing on YARN cluster with App id application_1707880073946_0002)

VERTICES	STATUS	TOTAL	COMPLETED	RUNNING	PENDING	FAILED	KILLED
Map 1	SUCCEEDED	3	3	0	0	3	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: 6.01 s

OK

1	Ross	25	Mumbai	25000	2024-02-14	NULL	NULL	NULL
2	Mike	27	Bhopal	35000	2024-02-14	101	2	20000
2	Mike	27	Bhopal	35000	2024-02-14	102	2	25000
3	Albin	24	Pune	50000	2024-02-14	NULL	NULL	NULL
NULL	NULL	NULL	NULL	NULL	NULL	104	4	10000

Time taken: 6.902 seconds, Fetched: 5 row(s)

hive> create view if not exists customers_vw as select * from customers where address="Mumbai";

OK

Time taken: 0.605 seconds

hive> select * from customers_vw;

OK

1	Ross	25	Mumbai	25000	2024-02-14
---	------	----	--------	-------	------------

Time taken: 0.271 seconds, Fetched: 1 row(s)

hive> alter view customers_vw as select * from customers;

OK

Time taken: 0.242 seconds

hive> select * from customers_vw;

OK

1	Ross	25	Mumbai	25000	2024-02-14
2	Mike	27	Bhopal	35000	2024-02-14
3	Albin	24	Pune	50000	2024-02-14

Time taken: 0.141 seconds, Fetched: 3 row(s)

hive> drop view if exists customers_vw;

OK

Time taken: 0.505 seconds

hive> select * from customers_vw;

FAILED: SemanticException [Error 10001]: Line 1:14 Table not found 'customers_vw'

hive> drop view if exists customers_vw; □