

Experiment 11:

Aim: Perform data processing operations using Hive

- a) Views
- b) Indexes

Entering into the hive environment and performing operations:

```
[cloudera@localhost ~]$ hive
Logging initialized using configuration in jar:file:/usr/lib/hive/lib/hive-common-0.10.0-cdh4.7.0.jar!/hive-log4j.properties
Hive history file=/tmp/cloudera/hive_job_log_b965e3bf-0fee-484a-a3b3-abbe937ba2b9_1567438058.txt
hive> desc student;
OK
sno      int
sname    string
m1       int
m2       int
m3       int
Time taken: 0.088 seconds
```

View:

Creating View:

Syntax: create view [if not exists] view_name [(c1,c2,...,cn)]
as select statement

```
hive> create view student_vw
> as select * from student;
OK
Time taken: 0.182 seconds
```

How to execute view:

Syntax: select * from view_name;

```
hive> select * from student_vw;
Total MapReduce jobs = 1
Launching Job 1 out of 1
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_202311091950_0001, Tracking URL = http://0.0.0.0:50030/jobdetails.jsp?jobid=job_202311091950_0001
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_202311091950_0001
```

Output:

```
Total MapReduce CPU Time Spent: 260 msec
OK
501    siri    10      20      30
502    giri    20      30      40
503    anu     30      40      50
510    anusha  50      10      44
506    pooja   44      20      66
517    sujatha 22      33      55
Time taken: 10.233 seconds
```

alter or change view:

Syntax: alter view[database_name] view_name as select statement;

Drop view:

Syntax: drop view [if exists] database_name;

```
hive> drop view if exists student_vw;  
OK  
Time taken: 0.5 seconds
```

Index:

Indexing are classified into 2 types.

- 1.Compact indexing
- 2.Bitmap indexing

Compact Indexing:

Creating:

Syntax:create index <index_name> on table <table_name><column_name>
as
'org.apache.hadoop.hive ql.index.compact.CompactIndexHandler'
WITH DEFERRED REBUILD;
hive> create index index_sno on table student(sno)
> as
> 'org.apache.hadoop.hive ql.index.compact.CompactIndexHandler'
> WITH DEFERRED REBUILD;

```
OK  
Time taken: 0.548 seconds
```

To see index details:

syntax:show index on table_name;

```
hive> show index on student;  
OK  
index_sno          student          sno          default_  
student_index_sno compact  
Time taken: 0.641 seconds
```

Alter:

Syntax:alter index index_name on table_name REBUILD;

Drop:

Syntax:drop index index_name on table_name;

```
hive> drop index index_sno on student;  
OK  
Time taken: 0.081 seconds
```

BITMAP indexing:

Creating:

Syntax:create index <index_name> on table <table_name><column_name>
as
'BITMAP'
WITH DEFERRED REBUILD;
hive> create index index_sno on table student(sno)
> as
> 'BITMAP'
> WITH DEFERRED REBUILD;

```
OK  
Time taken: 0.09 seconds
```

To see index details:

syntax:show index on table_name;