

Poornima Institute of Engineering & Technology
Department of Computer Engineering
Session 2020-21
IV year---VIII Semester
Subject: Big Data Analytics (8CS4-1)
Experiment-8
Topic: -Working with hive

Objective: Install and Run Hive then use Hive to create, alter, and drop databases, tables, views, functions, and indexes.

Hive is pre-installed in cloudera framework.

Step

1. Open Terminal – run command hive---enter

2. In Hive mode now

hive>

3. To create Database:-

hive> CREATE DATABASE user;

OK

Time taken: 4.478 seconds

4. Show database in hive:-

hive> show databases;

OK

default

user

Time taken: 0.77 seconds, Fetched: 2 row(s)

5. Drop Database:-

hive> DROP DATABASE IF EXISTS user;

OK

Time taken: 0.751 seconds

6. Create Table:-

hive> CREATE TABLE IF NOT EXISTS employee (eid int,name String,salary String,destination String)

> COMMENT 'Employee details'

> ROW FORMAT DELIMITED

> FIELDS TERMINATED BY '\t'

> LINES TERMINATED BY '\n'

> STORED AS TEXTFILE;

OK

Time taken: 2.316 seconds

7. Show Tables:-

hive>show table;

OK

emp

Time taken: 0.77 seconds, Fetched: 1 row(s)

8. Create txt file in /home/cloudera/hive/sample.txt.

+write data in file

9. Load data in emp table:

```
hive> LOAD DATA LOCAL INPATH '/home/cloudera/hive/sample.txt'
OVERWRITE INTO TABLE emp;
Loading data to table default.emp
Table default.emp stats: [numFiles=1, numRows=0, totalSize=201, rawDataSize=0]
OK
Time taken: 1.409 seconds
```

10. Hive - Alter Table

The statement takes any of the following syntaxes based on what attributes we wish to modify in a table.

```
ALTER TABLE name RENAME TO new_name
ALTER TABLE name ADD COLUMNS (col_spec[, col_spec ...])
ALTER TABLE name DROP [COLUMN] column_name
ALTER TABLE name CHANGE column_name new_name new_type
ALTER TABLE name REPLACE COLUMNS (col_spec[, col_spec ...])
```

11. The following query renames the table from employee to emp.

```
hive> ALTER TABLE employee RENAME TO emp;
OK
Time taken: 0.23 seconds
hive> show tables;
OK
emp
p
patients
prescription
Time taken: 0.05 seconds, Fetched: 4 row(s)
```

12. Change Statement

Change column name:-

```
hive> ALTER TABLE emp CHANGE name ename String;
OK
Time taken: 5.781 seconds
```

13. Change type of column:-

```
hive> ALTER TABLE emp CHANGE salary salary Double;
OK
Time taken: 0.269 seconds
```

14. To add new column in table:-

This query adds a column named dept to the emp table.

```
hive> ALTER TABLE emp ADD COLUMNS (dept STRING COMMENT 'Department
name');
OK
Time taken: 0.281 seconds
```

15. Replace Statement:

This query deletes all the columns from the emp table and replaces it with emp and name columns:

```
hive> ALTER TABLE emp REPLACE COLUMNS (eid INT empid Int,ename STRING  
name String);
```

16. Drop Table Statement:-

This query drops a table named employee:

```
hive> DROP TABLE IF EXISTS emp;
```

17. To create data randomly use www.generatedata.com, to see how it will work use my recorded video on teams.

https://poornimauniversity.sharepoint.com/sites/PIET_CS_VIIISEMSECTIONB/Shared%20Documents/BIG%20DATA%20ANALYTICS%20LAB/Recordings/Experiment-8-20210605_091115-Meeting%20Recording.mp4?web=1