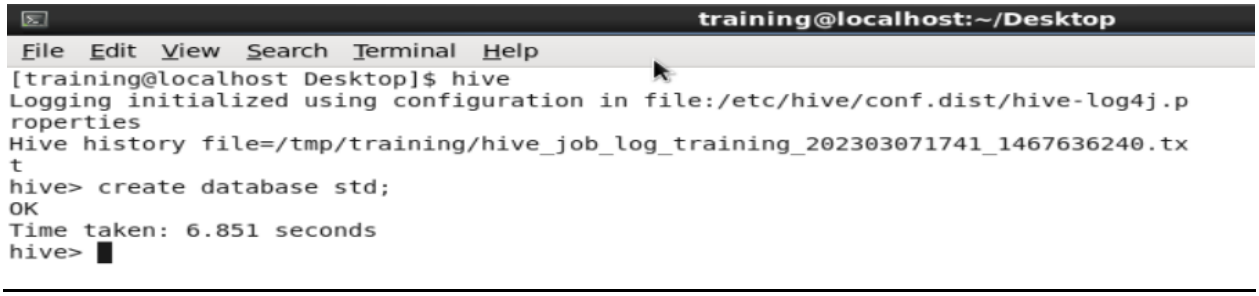
Name=Aditya Anal

HIVE TUTORIAL:

For Internal Table Using Local File System:

That means data file present in local file system

1.Create a databases:



1.2Create a Internal Table:

```
hive> create table student(name string,rollno int,dept string)
    > row format delimited
    > fields terminated by ',';
OK
Time taken: 0.78 seconds
```

1.3Show The Create Table:

```
hive> show tables;
OK
student
Time taken: 0.356 seconds
hive>
```

---

2. Load the data into internal table:

```
hive> load data local inpath '/home/training/allfile/student' into table student
    > ;
Copying data from file:/home/training/allfile/student
Copying file: file:/home/training/allfile/student
Loading data to table std.student
OK
Time taken: 0.949 seconds
hive>
```

2.1 to show table in hadoop

Database name is std.db.

Table name is student.

```
[training@localhost allfile]$ hadoop fs -ls  /user/hive/warehouse
Found 3 items
drwxr-xr-x   - training supergroup          0 2023-02-15 19:21 /user/hive/wareho
use/bd.db
drwxr-xr-x   - training supergroup          0 2023-02-16 09:14 /user/hive/wareho
use/genre
drwxr-xr-x   - training supergroup          0 2023-03-07 17:47 /user/hive/wareho
use/std.db
[training@localhost allfile]$ hadoop fs -ls  /user/hive/warehouse/std.db
Found 1 items
drwxr-xr-x   - training supergroup          0 2023-03-07 17:54 /user/hive/wareho
use/std.db/student
[training@localhost allfile]$
```

3. Display the content of the table:

```
hive> select *from student
    > ;
OK
aditya anal     23      cse
rahul sharma    22      me
vinay kumar     21      it
ajay singh      23      cse
        NULL    NULL
Time taken: 0.51 seconds
hive>
```

Attribute information:

```
hive> describe student
    > ;
OK
name    string
rollno  int
dept    string
Time taken: 0.27 seconds
hive>
```

CASE: Add Another data in student table

This data is add in student table

1.Load the data into internal table:

```
hive> load data local inpath '/home/training/allfile/student1' into table student
    > ;
Copying data from file:/home/training/allfile/student1
Copying file: file:/home/training/allfile/student1
Loading data to table std.student
OK
Time taken: 0.397 seconds
hive>
```

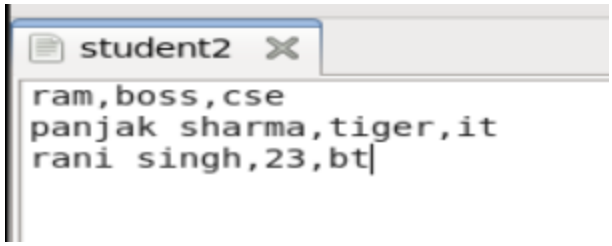
2. Display the Add content of the student table:

```
hive> select *from student
    > ;
OK
aditya anal     23      cse
rahul sharma    22      me
vinay kumar     21      it
ajay singh      23      cse
        NULL    NULL
sandeep saini   22      ce
anjali sharma   21      bt
rajeev ranjan   23      it
Time taken: 0.296 seconds
hive>
```

Case: to add different type of fields in student table :

To take a null value if fields type is not match

This type of data is add to student table



1. Load the data into internal table:

```
hive> load data local inpath '/home/training/allfile/student2' into table student
    > ;
Copying data from file:/home/training/allfile/student2
Copying file: file:/home/training/allfile/student2
Loading data to table std.student
OK
Time taken: 0.367 seconds
hive>
```

2. Display the content of the table:

```
hive> select * from student
    > ;
OK
aditya anal         23          cse
rahul sharma        22          me
vinay kumar         21          it
ajay singh          23          cse
          NULL      NULL
sandeep saini       22          ce
anjali sharma       21          bt
rajeev ranjan       23          it
ram         NULL    cse
panjak sharma       NULL        it
rani singh          23          bt
Time taken: 0.297 seconds
hive>
```

All datafile show in hdfs in hive path

```
[training@localhost allfile]$ hadoop fs -ls  /user/hive/warehouse/std.db
Found 1 items
drwxr-xr-x   - training supergroup          0 2023-03-07 18:19 /user/hive/warehouse/std.db/student
[training@localhost allfile]$ hadoop fs -ls  /user/hive/warehouse/std.db/student
Found 3 items
-rw-r--r--   1 training supergroup         75 2023-03-07 17:54 /user/hive/warehouse/std.db/student/student
-rw-r--r--   1 training supergroup         60 2023-03-07 18:10 /user/hive/warehouse/std.db/student/student1
-rw-r--r--   1 training supergroup         53 2023-03-07 18:19 /user/hive/warehouse/std.db/student/student2
[training@localhost allfile]$
```

For Internal Table Using HDFS File System:

That means data file present in HDFS file system

1.Create a file in hdfs:

```
[training@localhost allfile]$ hadoop fs -mkdir hivework1
[training@localhost allfile]$ hadoop fs -ls
Found 24 items
drwxr-xr-x   - training supergroup          0 2023-02-13 07:36 averageinput
drwxr-xr-x   - training supergroup          0 2023-02-13 07:40 averageoutput
drwxr-xr-x   - training supergroup          0 2023-02-11 11:48 distinctinput
drwxr-xr-x   - training supergroup          0 2023-02-11 11:50 distinctoutput
drwxr-xr-x   - training supergroup          0 2023-02-11 11:39 evenoddinput
drwxr-xr-x   - training supergroup          0 2023-02-11 11:41 evenoddoutput
drwxr-xr-x   - training supergroup          0 2023-02-16 08:22 hivework
drwxr-xr-x   - training supergroup          0 2023-03-07 18:30 hivework1
```

2.put data file into hdfs (file name is hivework1):

```
[training@localhost allfile]$ hadoop fs -put student hivework1
[training@localhost allfile]$ hadoop fs -put student1 hivework1
[training@localhost allfile]$ hadoop fs -put student2 hivework1
[training@localhost allfile]$ hadoop fs -ls hivework
Found 2 items
-rw-r--r--   1 training supergroup         41 2023-02-16 08:22 hivework/file.txt
-rw-r--r--   1 training supergroup        202 2023-02-14 11:28 hivework/u.genre
[training@localhost allfile]$ hadoop fs -ls hivework1
Found 3 items
-rw-r--r--   1 training supergroup         75 2023-03-07 18:34 hivework1/student
-rw-r--r--   1 training supergroup         60 2023-03-07 18:34 hivework1/student1
-rw-r--r--   1 training supergroup         53 2023-03-07 18:34 hivework1/student2
[training@localhost allfile]$
```

3.create table:

```
hive> use std
    > ;
OK
Time taken: 0.086 seconds
hive> create table student1(name string,rollno int,dept string)
    > row format delimited
    > fields terminated by ',';
OK
Time taken: 0.079 seconds
hive>
```

4.Load the data into internal table:

```
hive> load data  inpath 'hivework1/student' into table student1;
Loading data to table std.student1
OK
Time taken: 0.758 seconds
hive>
```

```
[training@localhost allfile]$ hadoop fs -ls  /user/hive/warehouse/std.db/student1
Found 1 items
-rw-r--r--   1 training supergroup         75 2023-03-07 18:34 /user/hive/warehouse/std.db/student1/student
[training@localhost allfile]$
```

4.Display the content of the table

```
hive> select *from student1
    > ;
OK
aditya anal     23      cse
rahul sharma    22      me
vinay kumar     21      it
ajay singh      23      cse
        NULL    NULL
Time taken: 0.584 seconds
hive>
```

External Table :

- External Table is loosely coupled in nature. Data will be available in HDFS.The table is going to create on HDFS data.
- In other way, we can say like its creating schema on data.
- At the time of dropping the table it drops only schema, the data will be still available in HDFS as before.
- External tables provide an option to create multiple schemas for the data stored in HDFS instead of deleting the data every time whenever schema updates

1.create a external table:

```
hive> create external table student2(name string,rollno int,dept string)
    > row format delimited
    > fields terminated by ','
    > location '/hivework1';
OK
Time taken: 0.205 seconds
hive>
```

2.Display the content of the table:

```
Time taken: 0.126 seconds
hive> select *from student1;
OK
aditya anal      23       cse
rahul sharma     22       me
vinay kumar      21       it
ajay singh       23       cse
          NULL     NULL
Time taken: 2.617 seconds
hive>
```