**Assignment 26.5**

How many kinds of tables are present in hive and explain the difference between them with a demo.

## Tables in Hive

As mentioned above, Hive has two types of tables:

* Managed table
* External table

## Managed table

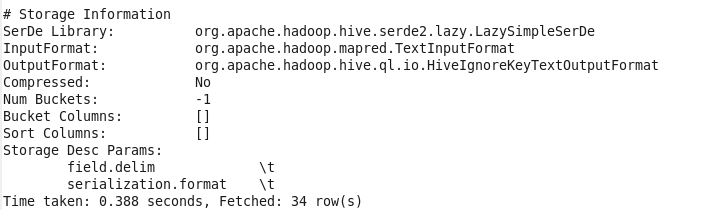
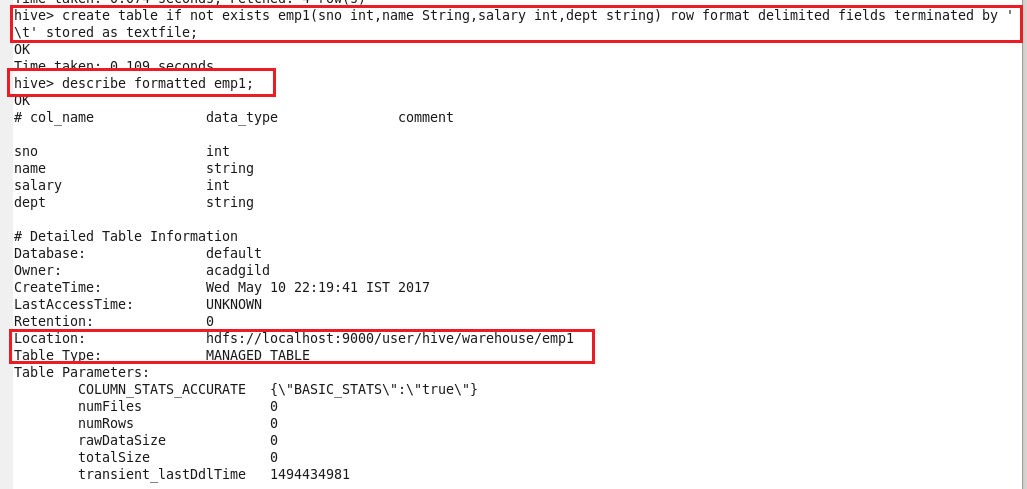
Managed table is also called as Internal table. This is the default table in Hive. When we create a table in Hive without specifying it as external, by default we will get a Managed table.

If we create a table as a managed table, the table will be created in a specific location in HDFS.By default, the table data will be created in **/usr/hive/warehouse**directory of HDFS.

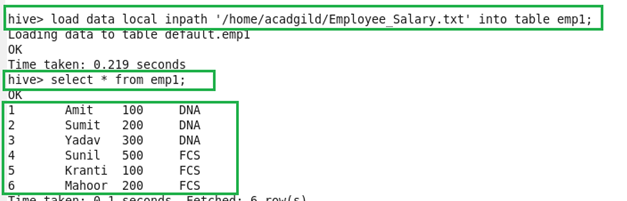
If we delete a Managed table, both the table data and meta data for that table will be deleted from the HDFS.

Let us create a managed table with the below command.

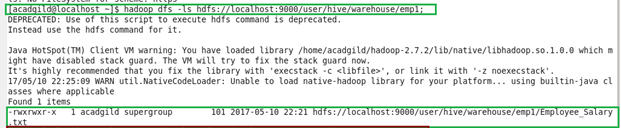
After creating the table to know the table type we use describe formatted ‘table name’ and it will show the type of the created table as we can see in the attached screenshot.



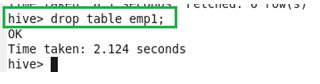
Now, loading the data into the created table. Using select keyword we can see the contents of table



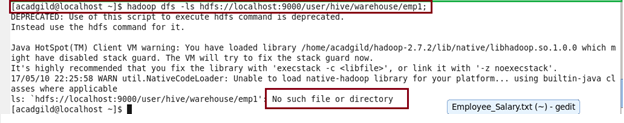
Now, checking the contents of the table in HDFS by using the below command.



Now, deleting the table using below command



Again, checking the contents of the table in the hdfs



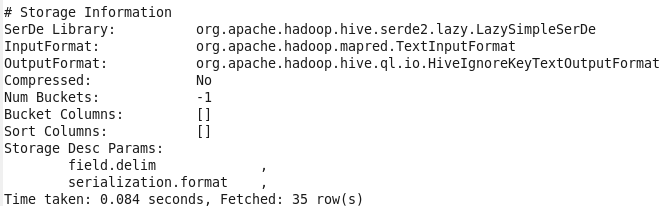
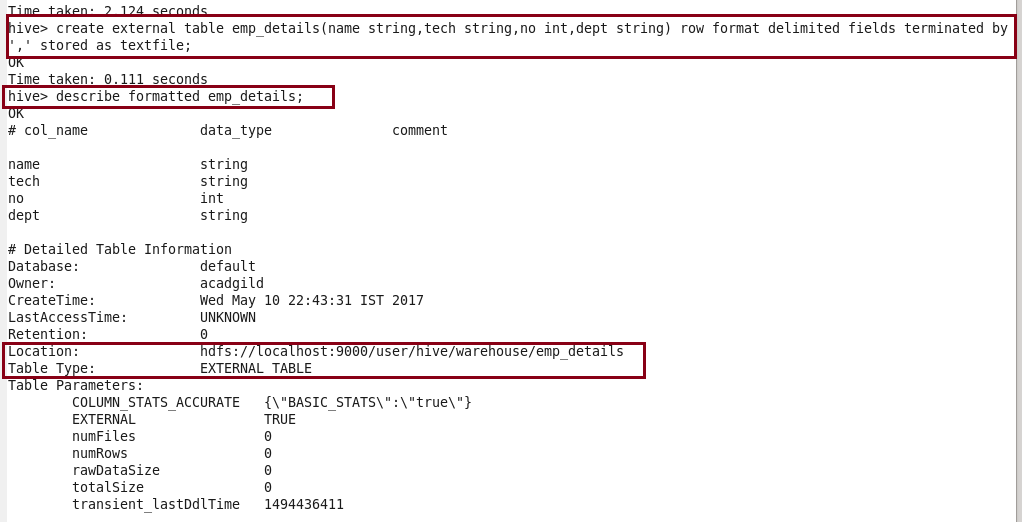
 It is displaying like **No such file or directory**because both the table and its contents are deleted from the HDFS location.

**EXTERNAL TABLE**

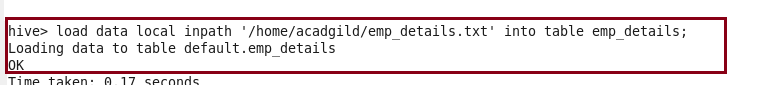
External table is created for external use as when the data is used outside Hive. Whenever we want to delete the table’s meta data and we want to keep the table’s data as it is, we use External table. External table only deletes the schema of the table.

Let us create an external table by using the below command.

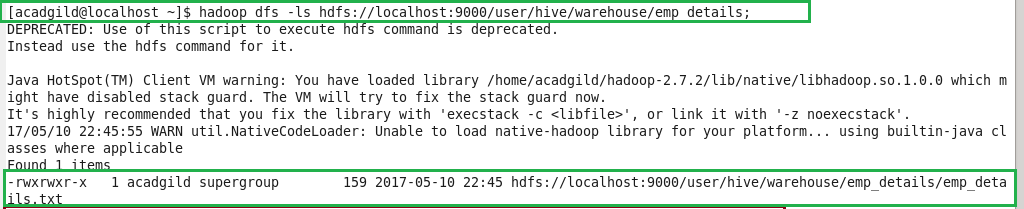
After creating the table to know the table type we use describe formatted ‘table name’ and it will show the type of the created table as we can see in the attached screenshot.



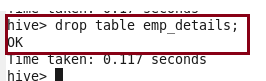
Now, loading the data into the created table. Using select keyword we can see the contents of table



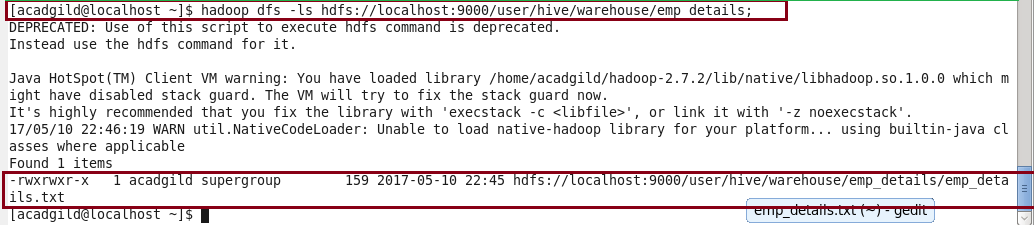
Now, checking the contents of the table in HDFS by using the below command.



Now, deleting the table using below command



Again, checking the contents of the table in the hdfs



It is displaying that the contents of the table are still present in the HDFS location.

If we create an External table, after deleting the table only the meta data related to table is deleted but not the contents of the table.

## When to use Managed table

* Data is temporary
* Hive to Manage the table data completely not allowing any external source to use the table
* Don’t want data after deletion

**When to use External table**

* The data is also used outside of Hive. For example, the data files are read and processed by an existing program that doesn’t lock the files
* Hive should not own data and control settings, dirs, etc., you have another program or process that will do those things
* You are not creating table based on existing table (AS SELECT)
* Can create table back and with the same schema and point the location of the data