**ASSIGNMENT26.5**

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

Hive has two types of tables:

* Managed table
* External table

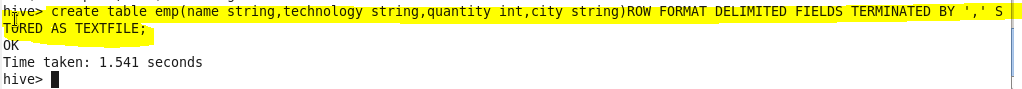
**Managed table**:

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 a Managed table is created.

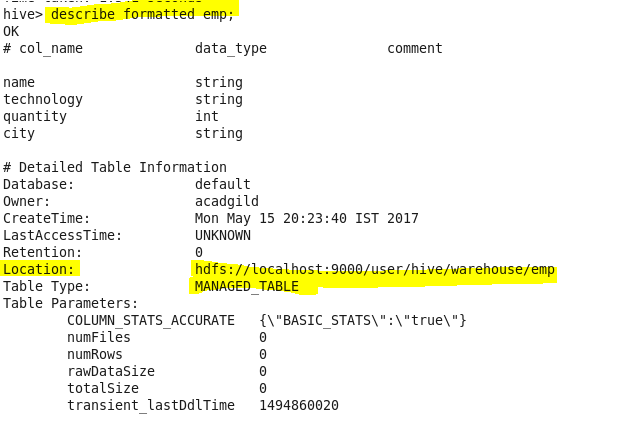
If we create a managed table, the table will be created in a specific location in HDFS.

By default, the table is created in **/user/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.



When we describe the table it shows it as a managed table.

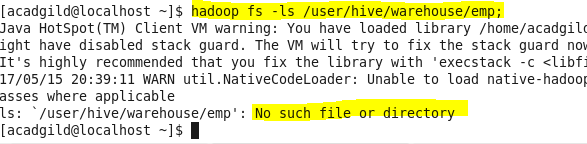


We will try to load one sample dataset which we have created into the table



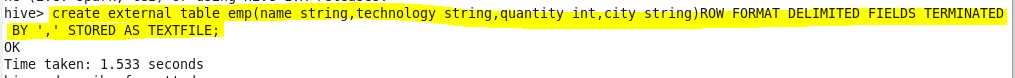
Now let us delete the above created table

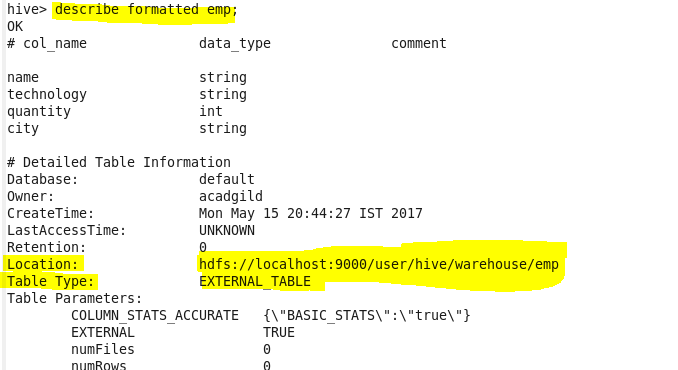




**EXTERNAL TABLE**

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.







Now let us delete the above created table



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.

The above approach will work only if your data is in **/user/hive/warehouse**directory. But if your data is in another location, if you delete the table the data will also get deleted. So in that case you need to mention the **external location**of the data while creating the table itself

## When to use External and Managed table

### 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

### 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