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

Hive have two kinds of tables. They are:

1. Managed or Internal tables
2. External tables

**Managed or Internal tables:**

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.

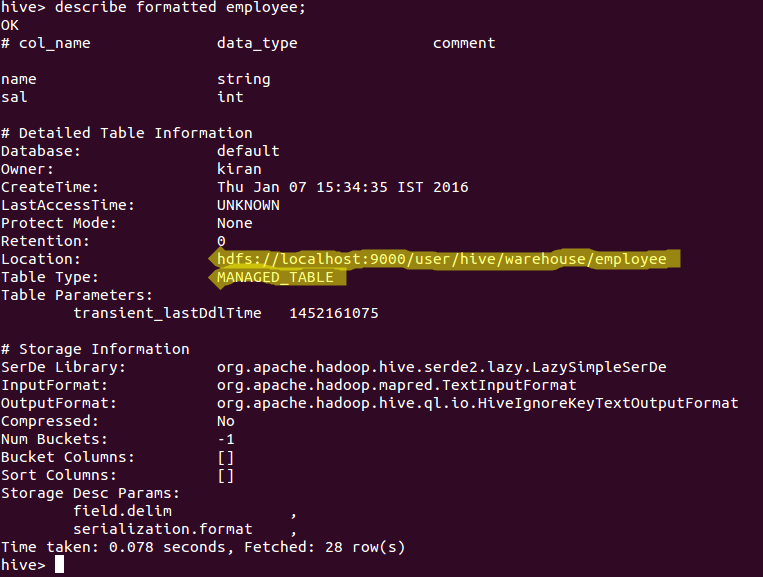
If we delete a Managed table, both the table data and meta data for that table will be deleted from the HDFS.

Let us create a managed table with the below command.

create table employee(Name String, Sal Int) row format delimited fields terminated by ',';

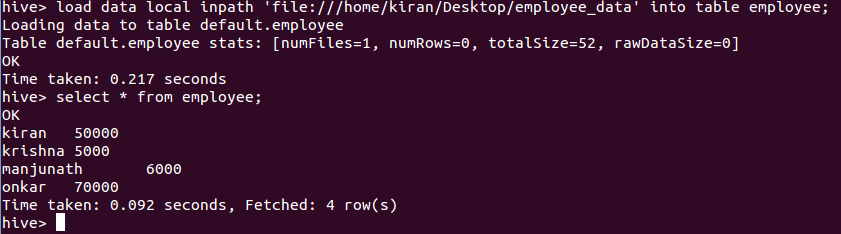
https://i2.wp.com/s3.amazonaws.com/acadgildsite/wordpress_images/bigdatadeveloper/Managed+and+External+tables+in+hive/Create_manage_table.png?resize=768%2C59&ssl=1We have successfully created the table and to check the details of the table type the below command:

describe formatted table\_name;

In the above image we can see**MANAGED\_TABLE as the entry for the option Table type**which means that we have created a Managed table.

We will try to load one sample dataset which we have created into the table by using the below command:

load data local inpath 'path of the file' into table employee;

If we check in the hdfs location we can get the contents of the table.

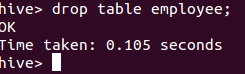
Check the contents of the table in HDFS by using the below command:

hadoop dfs -ls hdfs://localhost:9000/user/hive/warehouse/employee

https://i1.wp.com/s3.amazonaws.com/acadgildsite/wordpress_images/bigdatadeveloper/Managed+and+External+tables+in+hive/check_managed_data.png?resize=768%2C64&ssl=1In the above image we can see the contents of the table which is in the hdfs location.

Now let us delete the above created table by using the command

drop table employee;



We have successfully deleted the table.

Now let us try to check the contents of the table in HDFS using the below command:

hadoop dfs -ls hdfs://localhost:9000/user/hive/warehouse/employee

https://i0.wp.com/s3.amazonaws.com/acadgildsite/wordpress_images/bigdatadeveloper/Managed+and+External+tables+in+hive/check_data_after_delete_managed.png?resize=768%2C55&ssl=1In the above image, you can see that it is displaying like **No such file or directory**because both the table and its contents are deleted from the HDFS location.

Now let us create a table as External table.

**External tables:**

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.

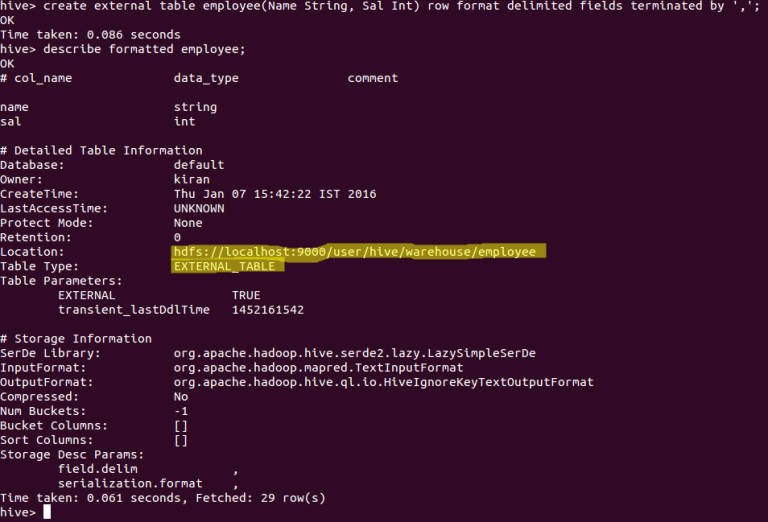
Let us create an external table by using the below command:

create external table employee(Name String, Sal Int) row format delimited fields terminated by ',';

We have now successfully created the external table.

Let us check the details regarding the table using the below command:

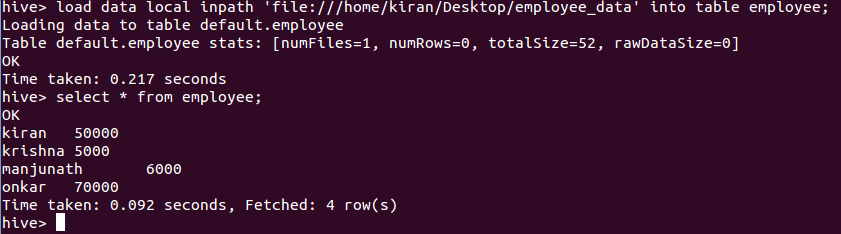
describe formatted employee;



In the above image we can see the **EXTERNAL\_TABLE as the entry for the option Table type** which says that the above table is an External table.

Now let us load some data into the table using the below command:

load data local inpath 'path of the file' into table employee;

We have successfully loaded data into the Hive table.

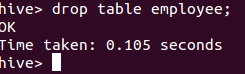
Let us check the contents in HDFS by using the below command:

hadoop dfs -ls hdfs://localhost:9000/user/hive/warehouse/employee

https://i1.wp.com/s3.amazonaws.com/acadgildsite/wordpress_images/bigdatadeveloper/Managed+and+External+tables+in+hive/check_managed_data.png?resize=768%2C64&ssl=1In the above image we can see the contents of the External table inside the HDFS.

Let’s now delete the table using the below command:

drop table employee;



We have successfully deleted the table.

Now let us check the HDFS location of the table using the below command:

hadoop dfs -ls hdfs://localhost:9000/user/hive/warehouse/employee

https://i0.wp.com/s3.amazonaws.com/acadgildsite/wordpress_images/bigdatadeveloper/Managed+and+External+tables+in+hive/check-data_after_delete_external.png?resize=768%2C72&ssl=1You can see 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.

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 as shown below.

create external table employee\_ext(Name String, Sal Int) row format delimited fields terminated by ',' LOCATION '/employee';

Here we have specified the location of the data in the table creation itself. Now if  you delete the table also the data will be there. But if you load the data explicitly using the **load**statement into the external table, if you drop the table now the data will also get deleted.