**ASSIGNMENT 26.5**

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

Two types of table are present and they are:

1. External table
2. Internal table

\***External Tables**: Hive will not move any data into the directory of the warehouse.

\*If there is a drop in the external table, then metadata of the table will be deleted, but not the data.

\***Internal tables**: Hive data will be moved into the directory of the warehouse.

\*If there is a drop in the internal table then both the metadata and the data will be deleted.

**External Tables:**

\*HDFS server is used to store the external table files and the tables

are not linked to any of the source file completely.

\*The file inside the HDFS will remain as it is even when the external table is deleted.

\***For example** – In HIVE when an external table is created it will be called as “table\_test”  by making use of HIVE-QL and the table will be linked to file by using “file”,

**\***Deleting table\_test will not effect in HDFS.It will remain in HDFS.

\*Anyone can access the external table files who has ability to access HDFS file structure

\*Master node will maintain the meta data and also for deleting any external table from HIVE, which only delete the contents of the metadata not any the data/file.

\***External tables are used when** the data is outside hive

**Internal Tables:**

\*Storage in directory is based on settings in hive.metastore.warehouse.dir internal tables are stored in the following directory

**“/user/hive/warehouse”**

\*One can change those by updating the location on the config file.

\*When a table is deleted the metadata and the data from the master-node as well from HDFS will be deleted respectively.

\*And the Security needs to be managed within HIVE, which means at the schema level (depending upon the organization).

\*Internal or external tables are present in HIVE and that is a choice which will affect how the data will be loaded, controlled, as well managed.

**INTERNAL tables are used when**:

The data is temporary.

The life-cycle of the table and data will be completely managed by Hive.