***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

* In case of ***External Tables*** ,Hive will not move any data into the directory of the warehouse.
* In case if there is a drop in the external table, then metadata of the table will be deleted
* but the data will not be deleted.
* For ***Internal tables*** , Hive data will be moved into the directory of the warehouse.
* In case if there is a drop in the external table then the table metadata will also be dropped.
* As well the data will also be deleted.

**For External Tables –**

* HDFS server was used to store the external table files instead the tables are not linked to any of the source file completely.
* When a external table is deleted then the file inside will remains on HDFS server.
* 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”,
* And fro deleting “table\_test” from the HIVE file will not delete “file” which is from the from HDFS.
* Anyone can access the external table files who has ability to access HDFS file structure
* For this purpose the security of the HDFS needs to be managed at file/folder level.
* 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.

**For Internal Tables-**

* Storage in directory will take place which is based on settings in hive.metastore.warehouse.dir,which will be by default 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.
* The control over internal table file security is through HIVE.
* 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.

**Use EXTERNAL tables when:**

* When the data used outside Hive.
* For example- when an existing program read and processed the data files that doesn’t lock the files.
* Then the data needs to be remain in the same location even after a DROP TABLE.
* This is applicable while pointing multiple schema (tables or views) at a single data set or else while one iterate through various schema.
* There will not be any data owned by Hive as well the control settings, directories, etc.,
* There were another program or process available for doing these type of process.
* One cannot create a table based on existing table (AS SELECT).
* Use INTERNAL tables are done only when:
* The data is temporary.
* The life-cycle of the table and data will be completely managed by Hive.