

⇒ In Databricks Lakehouse Platform, Schema and Database are same thing.

P.N: If we create a Schema, and check its location of creation using :

[DESCRIBE SCHEMA EXTENDED <schema-name>
in the Location field, we can see that the schema is an folder, created with .db extension

⇒ All tables, views, etc, created in / belonging to this schema, will have their files inside this .db folder.

⇒ Setting Location of Schema creation

CREATE SCHEMA <sch-name> LOCATION 'dbfs:/Folder/
<sch-name>.db'

we can set location of schema creation by using the LOCATION keyword, and mentioning the complete path (relative or absolute, along with schema/filename) after the keyword.

⇒ To use the created database/schema above, the command is :-

[USE <schema-name>

Physical view of data stored

in Delta Tables :-

⇒ In DBFS/file storage location, the objects created actually exist as follows :-



Created schema | when a delta table | inside each delta table
actually creates a | is created, actually a | folder, there is a
folder, and all the | folder is created with | folder named delta-log,
objects in it are | the same name, and | which records each
kept as files | all the data in | transaction in this
folders inside | records in this table | table as a separate
this folder. | are stored inside | JSON file.
 | this folder. |

— — — — — } And there are
 | parquet files, which
 | actually stores the data
 | in the table.

⇒ When a table is dropped, the table named folder and the log files inside it are deleted

⇒ DEFAULT Location of managed tables :-

⇒ If location is not specified while schema creation, & schema is not defined while table creation, managed table exist in following location :-

dbfs:/user/hive/warehouse/{schema-name.db}

⇒ CREATING EXTERNAL TABLE AT A CUSTOM LOCATION :-

We will create an external (unmanaged) table from sample data.

The data we are going to use are in CSV format. We want to create a Delta table with a **LOCATION** provided in the directory of our choice.

md 33

```
1 USE ${da.schema_name}_default_location;
2
3 CREATE OR REPLACE TEMPORARY VIEW temp_delays USING CSV OPTIONS (
4   path = '${da.paths.datasets}/flights/departuredelays.csv',
5   header = "true",
6   mode = "FAILFAST" -- abort file parsing with a RuntimeException if any malformed lines are encountered
7 );
8 CREATE OR REPLACE TABLE external_table LOCATION '${da.paths.working_dir}/external_table' AS
9   SELECT * FROM temp_delays;
10
11 SELECT * FROM external_table;
```