**Class 1 – Topics**

* Create Database
* Tables
* Internal vs External Tables
* Insert Statement
* Multi insert statement
* Alter table Schema
* Sorting Data

Check the mail for class 1 repositories

**CREATE DATABASES**

create database d1;

create database if not exists d1;

describe database;

**ADDING COMMENTS TO DATABASE**

create database if not exists d2 comment 'this is a database';

describe database extended d2;

**ADDING DB PROPERTIES TO DATABASE**

create database if not exists d3 ('creator'='jivesh','date'='2017-12-01');

describe database extended d3;

**FINDING DATABASES AND USING IT**

Show databases;

use d2;

**TABLE CREATION AND LOADING DATA INTO IT**

**Check the default warehouse location:** SET HIVE.METASTORE.WAREHOUSE.DIR

create table if not exists table3(col1 int,col2 array<string>,col3 string,col4 int)row format delimited fields terminated by',' collection items terminated by':' lines terminated by'\n' stored as textfile;

create external table if not exists table2(col1 int,col2 array<string>,col3 string,col4 int)row format delimited fields terminated by',' collection items terminated by':' lines terminated by'\n' stored as textfile location'/user/…/emp';

load data local inpath'/home/…/table1'into table table2;

**INTERNAL VS EXTERNAL TABLES**

**External table creation:**

CREATE EXTERNAL TABLE if not exists table table1(col1 string, col2 array<string>,col3 string, col4 string)row format delimited fields terminated by ',' collection items terminated by ':' lines terminated by '\n' stored as textfile LOCATION ‘/home…’;

lOAD data local inpath '/home....' into table table1;

DROP the table and see the results

External table’s data can be accessed by any application

Internal table’s data can be accessed only by hive

**INSERT STATEMENT :**

Moving data between tables

create table table2(id int,country string)stored as textfile ;

insert into table table2 select col1,col2,col3 from table1 ;

insert into table table3 using where condition

insert overwrite table table 3

**MULTI -INSERT STATEMENT:**

create table movies\_genre(name string,actors array<string>,genre string,rating int)row format delimited fields terminated by','collection items terminated by':' lines terminated by'\n' stored as textfile;

load data local inpath'/home/jivesh/files/movies\_genre'into table movies\_genre;

create table drama(name string,actors array<string>,genre string,rating int) stored as textfile;

create table fiction(name string,actors array<string>,genre string,rating int) stored as textfile;

from movies\_genre insert into table drama select name,actors,genre,rating where genre='drama' insert overwrite table fiction select name,actors,genre,rating where genre='science fiction';

**MOVIES\_GENRE :**

titanic,leonardo dicaprio:kate winslet,drama,7.7

avatar,zoe saldana:Sam worthington,science fiction,7.8

logan,hugh jackman:Patrick stewart,science fiction,8.4

forest gump,tom janks:robin wright,drama,8.8

**MOVIES:**

titanic,leonardo dicaprio:kate winslet:Billy zane,drama,7.7

avatar,zoe saldana:Sam worthington,science fiction,7.8

logan,hugh jackman:Patrick stewart,science fiction,8.4

**ALTER TABEL SCHEMA:**

create table if not exists table9(col1 int,col2 string,col3 string,col4 int)row format delimited fields terminated by',' lines terminated by'\n' stored as textfile;

load data local inpath'/home/jivesh/files/dynamic'into table table9;

alter table table5 rename to table10;

desc table9;

alter table table10 change column col3 expalination int;

alter table table10 change column expalination col3 string;

alter table table10 change column col1 id int after col2;

alter table table10 add columns (col5 int,col6 string);

desc table10;

select \* from table10;

alter table table10 replace columns (col7 int,col8 int);

desc table10;

select \* from table10;

alter table table10 set tblproperties('creator'='jivesh');

desc formatted table10;

alter table table10 set fileformat sequencefile;

**Questions:**

* How to change the default meta store ?
* How to use overwrite in load table command and what is the use of it?
* what is the port number of the file browser ?
* Where will my data gets stored if i dont provide default location in external table ?

**Assignment for Students**

1. Study the structure of data then use create statement to create a table with necessary number of columns.

2. Load the data into table.

3. View data in tables after loading.

You can also create 1 more table and practice Insert command by loading data into it from previous table.

Questions for this Assignment

Can I create a table with less number of columns than dataset?

Download Resource Files