1. Explain with an example in brief.

● Hive Data Definitions

● Hive Data Manipulations

● HiveQL Manipulations

 DDL statements are documented here, including:

* CREATE DATABASE/SCHEMA, TABLE, VIEW, FUNCTION, INDEX
* DROP DATABASE/SCHEMA, TABLE, VIEW, INDEX
* TRUNCATE TABLE
* ALTER DATABASE/SCHEMA, TABLE, VIEW
* MSCK REPAIR TABLE (or ALTER TABLE RECOVER PARTITIONS)
* SHOW DATABASES/SCHEMAS, TABLES, TBLPROPERTIES, VIEWS, PARTITIONS, FUNCTIONS, INDEX[ES], COLUMNS, CREATE TABLE
* DESCRIBE DATABASE/SCHEMA, table\_name, view\_name

Create Database

CREATE (DATABASE|SCHEMA) [IF NOT EXISTS] database\_name

[COMMENT database\_comment]

[LOCATION hdfs\_path]

[WITH DBPROPERTIES (property\_name=property\_value, ...)];

Drop Database

DROP (DATABASE|SCHEMA) [IF EXISTS] database\_name [RESTRICT|CASCADE];

Alter Database

ALTER (DATABASE|SCHEMA) database\_name SET DBPROPERTIES (property\_name=property\_value, ...);

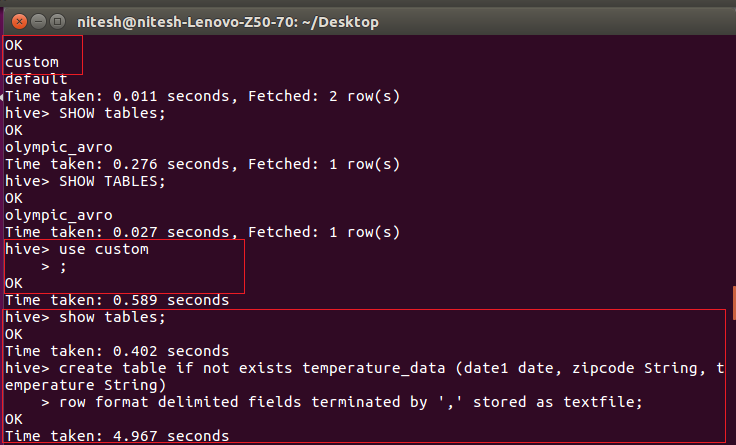
ALTER (DATABASE|SCHEMA) database\_name SET OWNER [USER|ROLE] user\_or\_role;

Use Database

USE database\_name;

USE DEFAULT;

C:\Users\612840.CTS\Desktop\sub_screen_shots\26_3\23_3_part_table.JPG



CREATE TABLE creates a table with the given name. An error is thrown if a table or view with the same name already exists. You can use IF NOT EXISTS to skip the error.

* Table names and column names are case insensitive but SerDe and property names are case sensitive.
  + In Hive 0.12 and earlier, only alphanumeric and underscore characters are allowed in table and column names.
  + In Hive 0.13 and later, column names can contain any [Unicode](http://en.wikipedia.org/wiki/List_of_Unicode_characters) character, however, dot (**.**) and colon (**:**) yield errors on querying, so they are disallowed in Hive 1.2.0. Any column name that is specified within backticks (`) is treated literally. Within a backtick string, use double backticks (``) to represent a backtick character. Backtick quotation also enables the use of reserved keywords for table and column identifiers.

1) Hive Data Definitions

* The structures of table in a database are included in the Hive Data Definitions
* Hive Data Definition language is the derived set of Hive SQL statements, this describes the structure of the data in the table creating, deleting, or altering schema objects such as databases, tables, views are done in HDD.

The commands that are under HDD:

CREATE, DROP, TRUNCATE, DESCRIBE, AND ALTER.

Syntax:

CREATE TABLE HIVE\_TABLE (abc INT, xyz STRING); // creating tables

DROP ‘TABLE NAME’; // deleting the table or database

TRUNCATE – used to delete the structure of the table.

ALTER - used to change the structure of the table.

2) Hive Data Manipulations

Hive Data Manipulations are used to modify the values in the table or to extract the data from the table.

The commands are:

LOAD, INSERT, UPDATE, DELETE

-LOAD :- load the data into a table.

-Update :- change the entries in the table.

3) HiveQL Manipulations

HiveQL Manipulations are of three types:

a) Select where

b) Select order by /sort by

c) Select group by

Select where :- used to filter the dataset with the where clause and apply the conditions.

Select order by clause :- used to arrange the dataset in the ascending or descending order by one field or multiple fields.

Select group by :- used to form subsets of the database using different fields of the data.