**Assignment 7.3:**

**Problem Statement**:

Explain the below concepts

* Hive Data Definitions
* Hive Data Manipulations
* HiveQL Manipulations with an example in brief.
* **Hive Data Definitions:**

Data Definition Language (DDL)

DDL statements are used to build and modify the tables and other objects in the database.

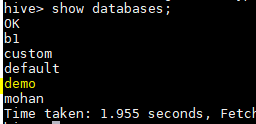
Example :

CREATE, DROP, TRUNCATE, ALTER, SHOW, DESCRIBE Statements.

Suppose to create a database **demo** the hive command is :

hive>CREATE database demo;

hive>show databases;



* **Hive Data Manipulations**

Data Manipulation Language (DML )

DML statements are used to retrieve, store, modify, delete, insert and update data in the database.

Example :

LOAD, INSERT Statements.

Syntax :

LOAD data <LOCAL> inpath <file path> into table [tablename]

The Load operation is used to move the data into corresponding Hive table. If the keyword local is specified, then in the load command will give the local file system path. If the keyword local is not specified we have to use the HDFS path of the file.

Suppose to load data to **Employee** table of demo database, below hive DML command is to be used:

Hive> LOAD DATA LOCAL INPATH '/home/acadgild/hive/emp\_details.txt'

INTO TABLE demo.Employee;

* **HiveQL Manipulations:**

Aggregation :

Select count (DISTINCT category) from tablename;

This command will count the different category of ‘cate’ table. Here there are 3 different categories.

Suppose to find number of employee having different skills, we need to use below hive query:

SELECT COUNT(\*) FROM demo.Employee GROUP BY skill;

