Assignment-24.1:

Explain with an example in brief.

● Hive Data Definitions

● Hive Data Manipulations

● HiveQL Manipulations

**\* Hive Data Definitions:**

HDL is a subset of Hive SQL statements that describe the data structure in Hive by creating, deleting, or altering schema objects such as databases, tables, views, partitions, and buckets.

**CREATE - to create objects in the database**

Syntax: create table table\_name (id int, name string);

**ALTER - alters the structure of the database**

Syntax: alter table table\_name add column column\_name;

**DROP - delete objects from the database**

Syntax: drop table table\_name;

**TRUNCATE - remove all records from a table, including all spaces allocated for the records are removed**

Syntax: truncate table table\_name;

\* **Hive Data Manipulations:**

Hive data manipulations commands are used for modifying the values in the table or extracting data from the table.

E.g. LOAD, INSERT, UPDATE, DELETE

**LOAD: loads the data set into the table**

Syntax: Load data local inpath ‘/path/’ into table\_name;

**INSERT: inserting values int the table**

Syntax: Insert into table table\_name values(…….);

**UPDATE: updates the rows in the table**

Syntax: UPDATE table\_name  
SET column1 = value1, column2 = value2, ...  
WHERE condition;

**DELETE: deletes the rows in the table**

Syntax: DELETE FROM table\_name  
WHERE condition;

**\* HiveQL Manipulations**

Hiveql manipulations are of three types

1. Select where
2. Select order by /sort by
3. Select group by

Select where is used to filter the dataset with the where clause and apply conditions. SELECT statement is used to retrieve the data from a table. WHERE clause works similar to a condition. It filters the data using the condition and gives you a finite result. The built-in operators and functions generate an expression, which fulfils the condition.

Select \* from table\_name where id=1;

The select order by clause is used to arrange the dataset in the ascending and descending order by one field or multiple fields.

Select \* from table order by id;

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

Select dept,count(id) from table\_name group by dept