**Problem Statement:**

**Explain the below concepts with an example in brief.**

**● Hive Data Definitions :**

Hive data definiation language is subset of Hive SQL statements that describes the data structure in Hive by creating , or altering schema object such as databases, tables, views, partitions, and buckets. Most of hive DDL statements start with keywords, **CREATE , DROP or ALTER**

HiveQL 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,

The simplest syntax for creating a database is shown in the following example

**CREATE DATABASE financials**

**Or**

**CREATE DATABASE IF NOT EXISTS financials;**

create table Olympics (Athlete string , Age string,Country string, Year int , Closing\_Date string,Sport string,Gold\_Medals int,Silver\_Medals int,Bronze\_Medals int,Total\_Medals int) comment 'Olympics details'

Row format delimited

fields terminated by ', '

stored as textfile location '/home/acadgild/olympix\_data.xls';

**● Hive Data Manipulations**

data manipulation language parts that are used to put data into tables and to extract data from tables to the filesystem.

e.g.

Load data local inpath ‘/home/acadgild/olympix\_data.csv’ into table Olympics;

**● HiveQL Manipulations**

Hive defines a simple SQL-like query language to querying and managing large datasets called Hive-QL ( HQL ). It’s easy to use if you’re familiar with SQL Language. Hive allows programmers who are familiar with the language to write the custom MapReduce framework to perform more sophisticated analysis.

**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.